99/19743

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/54, 9/12, C12Q 1/68, C07K 16/40, C12Q 1/48, G01N 33/50, A61K 35/00, C12N 15/11

A2

(11) International Publication Number: WO 99/66051

(43) International Publication Date: 23 December 1999 (23.12.99)

(21) International Application Number: PCT/US99/13533

(22) International Filing Date: 15 June 1999 (15.06.99)

(30) Priority Data:

60/089,520	16 June 1998 (16.06.98)	US
60/098,265	28 August 1998 (28.08.98)	US

(71) Applicant *(for all designated States except US)*: SUGEN, INC. [US/US]; 230 East Grand Avenue, South San Francisco, CA 94080 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: PLOWMAN, Gregory [US/US]; 4 Honeysuckle Lane, San Carlos, CA 94070 (US). MARTINEZ, Ricardo [US/US]; 984 Cartier Lane, Foster City, CA 94404 (US). ZHU, Yingfang [US/US]; 152 Oxford Lane #1, San Bruno, CA 94066 (US).

(74) Agent: WARBURG, Richard, J.; Lyon & Lyon LLP, Suite 4700, 633 West Fifth Street, Los Angeles, CA 90071-2066 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: NEK-RELATED AND BUB1-RELATED PROTEIN KINASES

(57) Abstract

The present invention relates to BUB1 and NEK kinase polypeptides, nucleotide sequences encoding the kinase polypeptides, as well as various products and methods useful for the diagnosis and treatment of various kinase-related diseases and conditions. Through the use of a targeted PCR cloning strategy, the human BUB1 serine/threonine kinase was identified and its protein structure was predicted. Additionally, through the use of a bioinformatics strategy, three mammalian members of the NEK-subfamily of STK's have been identified and their protein structure predicted.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

DESCRIPTION

NEK-RELATED AND BUB1-RELATED PROTEIN KINASES

FIELD OF THE INVENTION

The present invention relates to novel kinase polypeptides, nucleotide sequences encoding the novel kinase polypeptides, as well as various products and methods useful for the diagnosis and treatment of various kinase-related diseases and conditions.

BACKGROUND OF THE INVENTION

The following description of the background of the invention is provided to aid in understanding the invention, but is not admitted to be or to describe prior art to the invention.

Cellular signal transduction is a fundamental mechanism whereby external stimuli that regulate diverse cellular processes are relayed to the interior of cells. One of the key biochemical mechanisms of signal transduction involves the reversible phosphorylation of proteins, which enables regulation of the activity of mature proteins by altering their structure and function.

The best characterized protein kinases in eukaryotes phosphorylate proteins on the hydroxyl substituent of serine, threonine and tyrosine residues. These kinases largely fall into two groups, those specific for phosphorylating serines and threonines, and those specific for phosphorylating tyrosines. Some kinases, referred to as "dual specificity" kinases, are able to phosphorylate on tyrosine as well as serine/threonine residues.

Protein kinases can also be characterized by their location within the cell. Some kinases are transmembrane receptor-type proteins capable of directly altering their catalytic activity in response to the external environment such as the binding of a ligand. Others are non-receptor-type proteins lacking any transmembrane domain. They can be found in a variety of cellular compartments from the inner surface of the cell membrane to the nucleus.

Many kinases are involved in regulatory cascades wherein their substrates may include other kinases whose activities are regulated by their phosphorylation state. Ultimately the activity of some downstream effector is modulated by phosphorylation resulting from activation of such a pathway.

Protein kinases are one of the largest families of eukaryotic proteins with several hundred known members. These proteins share a 250-300 amino acid domain that can be subdivided into 12 distinct subdomains that comprise the common catalytic core structure. These conserved protein motifs have recently been exploited using PCR-based cloning strategies leading to a significant expansion of the known kinases.

Multiple alignment of the sequences in the catalytic domain of protein kinases and subsequent parsimony analysis permits the segregation of related kinases into distinct branches of subfamilies including: tyrosine kinases, dual-specificity kinases, and serine/threonine kinases (STK's). The latter subfamily includes cyclic-nucleotide-dependent kinases, calcium/calmodulin kinases, cyclin-dependent kinases (CDK's), MAP-kinases, serine-threonine kinase receptors, and several other less defined subfamilies

Cell cycle checkpoints are regulatory pathways that control the order and timing of cell cycle transitions and ensure that critical events such as DNA replication and chromosome segregation are completed with high fidelity. The mitotic checkpoint ensures proper chromosome segregation by delaying anaphase until all chromosomes are aligned on the spindles. Loss of cell cycle checkpoints in mitosis can result in cell death or uncontrolled cell division.

Checkpoint losses, resulting in genomic instability, have been implicated in the evolution of normal cells into cancer cells. Cancer cells differ from normal cells in many important characteristics including differentiation, increased invasiveness and decreased drug sensitivity.

CDK's play a pivotal role in the regulation of mitosis during cell division. The process of cell division occurs in four stages: S phase, the period during which

chromosomes duplicate; G2; mitosis; and G1, or interphase. During mitosis the duplicated chromosomes are evenly segregated, allowing each daughter cell to receive a complete copy of the genome. A key mitotic regulator in all eukaryotic cells is the STK cdc2, a CDK regulated by cyclin B.

SUMMARY OF THE INVENTION

Novel serine/threonine kinases (STK's) have been identified and their protein sequence predicted by as part of the instant invention. Mammalian members of the NEK family were identified through the use of a bioinformatics strategy, and the human BUB1 kinase was identified through the use of a targeted PCR cloning strategy. The partial or complete sequences of these kinases are presented here, together with their classification, predicted or deduced protein structure, and an assessment of their pattern of expression.

The NEK kinases of the invention function in cell cycle regulation. Consequently, the deregulation of these kinases through overexpression and/or mutational events plays a role in cancer. Similarly, they also play a role in neurodegenerative and/or immune disorders through aberrant modulation of cell division, differentiation, and apoptosis. Thus, modulators of NEK kinases will play a role in the treatment and/or prevention of cancers and neurodegenerative or immune disorders in which a NEK kinase is functioning aberrantly through overexpression or as the result of a mutational event.

In addition to their use as clinical targets, a better understanding of the role of NEK kinases in cell cycle regulation will have an impact on issues as diverse as cellular differentiation and apoptosis, a better understanding of which will impact things as diverse as aging, nerve regeneration, and developmental abnormalities. Various claimed aspects of this invention, including inhibitors, probes, antibodies, and hybridomas to the proteins or genes of the invention, will facilitate these studies.

The human *bub1* gene was found to be a checkpoint gene in mitosis. The biological and clinical therapeutic relevance of BUB1 was demonstrated by expression analysis which showed that the *bub1* transcript was expressed at very low levels in normal human tissues, but was strikingly overexpressed in numerous tumor cell lines. The high expression of the *bub1* gene in many malignant tumor cells, indicates that deregulation of BUB1 may play a role in the development of cancer. Overexpression of the *bub1* gene in cancer cells also suggests that it may be an attractive target for cancer therapy.

BUB1 expression was found to be cell cycle regulated and located in kinetochores, indicating that BUB1 is a centromeric protein involved in chromosome segregation during mitosis. In addition to its potential role in tumorigenesis, BUB1 may also be the target of human autoimmune diseases such as CREST syndrome typified by the presence of auto-antibodies to centromeric proteins.

Thus, a first aspect of the invention features an isolated, enriched, or purified nucleic acid molecule encoding a kinase polypeptide selected from the group consisting of a BUB1 kinase having the amino acid sequence set forth in SEQ ID NO:2, a NEK4a kinase having the amino acid sequence set forth in SEQ ID NO:8, a NEK4b kinase having the amino acid sequence set forth in SEQ ID NO:10, a NEK5 kinase having the amino acid sequence set forth in SEQ ID NO:4, and a NEK6 kinase having the amino acid sequence set forth in SEQ ID NO:6.

By "isolated" in reference to nucleic acid is meant a polymer of 6 (preferably 21, more preferably 39, most preferably 75) or more nucleotides conjugated to each other, including DNA and RNA that is isolated from a natural source or that is synthesized. In certain embodiments of the invention, longer nucleic acids are preferred, for example those of 300, 600, 900 or more nucleotides and/or those having at least 50%, 60%, 75%, 90%, 95% or 99% identity to the sequence shown in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO 5, SEQ ID NO:7, or SEQ ID NO:9.

The isolated nucleic acid of the present invention is unique in the sense that it is not found in a pure or separated state in nature. Use of the term "isolated" indicates that a naturally occurring sequence has been removed from its normal cellular (*i.e.*, chromosomal) environment. Thus, the sequence may be in a cell-free solution or placed in a different cellular environment. The term does not imply that the sequence is the only nucleotide chain present, but that it is essentially free (about 90 - 95% pure at least) of non-nucleotide material naturally associated with it, and thus is distinguished from isolated chromosomes.

By the use of the term "enriched" in reference to nucleic acid is meant that the specific DNA or RNA sequence constitutes a significantly higher fraction (2- to 5-fold) of the total DNA or RNA present in the cells or solution of interest than in normal or diseased cells or in the cells from which the sequence was taken. This could be caused by a person by preferential reduction in the amount of other DNA or RNA present, or by a preferential increase in the amount of the specific DNA or RNA sequence, or by a

combination of the two. However, it should be noted that enriched does not imply that there are no other DNA or RNA sequences present, just that the relative amount of the sequence of interest has been significantly increased. The term "significant" is used to indicate that the level of increase is useful to the person making such an increase, and generally means an increase relative to other nucleic acids of about at least 2-fold, more preferably at least 5- to 10-fold or even more. The term also does not imply that there is no DNA or RNA from other sources. The other source DNA may, for example, comprise DNA from a yeast or bacterial genome, or a cloning vector such as pUC19. This term distinguishes from naturally occurring events, such as viral infection, or tumor type growths, in which the level of one mRNA may be naturally increased relative to other species of mRNA. That is, the term is meant to cover only those situations in which a person has intervened to elevate the proportion of the desired nucleic acid.

It is also advantageous for some purposes that a nucleotide sequence be in purified form. The term "purified" in reference to nucleic acid does not require absolute purity (such as a homogeneous preparation). Instead, it represents an indication that the sequence is relatively more pure than in the natural environment (compared to the natural level this level should be at least 2- to 5-fold greater, *e.g.*, in terms of mg/mL). Individual clones isolated from a cDNA library may be purified to electrophoretic homogeneity. The claimed DNA molecules obtained from these clones could be obtained directly from total DNA or from total RNA. The cDNA clones are not naturally occurring, but rather are preferably obtained via manipulation of a partially purified naturally occurring substance (messenger RNA). The construction of a cDNA library from mRNA involves the creation of a synthetic substance (cDNA) and pure individual cDNA clones can be isolated from the synthetic library by clonal selection of the cells carrying the cDNA library. Thus, the process which includes the construction of a cDNA library from mRNA and isolation of distinct cDNA clones yields an approximately 10^6-fold purification of the native message. Thus, purification of at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated.

By a "kinase polypeptide" is meant 32 (preferably 40, more preferably 45, most preferably 55) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:2, 33 (preferably 40, more preferably 45, most preferably 55) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:4, or 8 (preferably 15, more preferably 20, most preferably 30) or more contiguous amino acids

set forth in the amino acid sequence of SEQ ID NO:6, 33 (preferably 40, more preferably 45, most preferably 55) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:8, 33 (preferably 40, more preferably 45, most preferably 55) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:10, or functional derivatives thereof as described herein. In certain aspects, polypeptides of 100, 200, 300 or more amino acids are preferred. The kinase polypeptide can be encoded by a full-length nucleic acid sequence or any portion of the full-length nucleic acid sequence, so long as a functional activity of the polypeptide is retained. It is well known in the art that due to the degeneracy of the genetic code numerous different nucleic acid sequences can code for the same amino acid sequence. Equally, it is also well known in the art that conservative changes in amino acid can be made to arrive at a protein or polypeptide which retains the functionality of the original. In both cases, all permutations are intended to be covered by this disclosure.

The amino acid sequence will be substantially similar to the sequence shown in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10, or the corresponding full-length amino acid sequence, or fragments thereof.

A sequence that is substantially similar to the sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10 will preferably have at least 90% identity (more preferably at least 95% and most preferably 99-100%) to the sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10.

By "identity" is meant a property of sequences that measures their similarity or relationship. Identity is measured by dividing the number of identical residues by the total number of residues and gaps and multiplying the product by 100. "Gaps" are spaces in an alignment that are the result of additions or deletions of amino acids. Thus, two copies of exactly the same sequence have 100% identity, but sequences that are less highly conserved, and have deletions, additions, or replacements, may have a lower degree of identity. Those skilled in the art will recognize that several computer programs are available for determining sequence identity using standard parameters, for example Gapped BLAST or PSI-BLAST (Altschul, *et al.* (1997) Nucleic Acids Res. 25:3389-3402), BLAST (Altschul, *et al.* (1990) J. Mol. Biol. 215:403-410), and Smith-Waterman (Smith, *et al.* (1981) J. Mol. Biol. 147:195-197).

In preferred embodiments, the invention features isolated, enriched, or purified nucleic acid molecules encoding a kinase polypeptide comprising a nucleotide sequence that: (a) encodes a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10; (b) is the complement of the nucleotide sequence of (a); (c) hybridizes under highly stringent conditions to the nucleotide molecule of (a) and encodes a naturally occurring kinase polypeptide; (d) encodes a kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10, except that it lacks one or more, but not all, of the following segments of amino acid residues: 1-785, 786-1028, or 1029-1085 of SEQ ID NO:2, 1-31, 32-283, or 284-302 of SEQ ID NO:4, 1-22 or 23-302 of SEQ ID NO:6, (e) is the complement of the nucleotide sequence of (d); (f) encodes a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 from amino acid residues 1-785, 786-1028, or 1029-1085 of SEQ ID NO:2, 1-31, 32-283, or 284-302 of SEQ ID NO:4, 1-22 or 23-302 of SEQ ID NO:6, (g) is the complement of the nucleotide sequence of (f); (h) encodes a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 SEQ ID NO:8, or SEQ ID NO:10, except that it lacks one or more, but not all, of the domains selected from the group consisting of a N-terminal domain, a catalytic domain, a C-terminal catalytic domain, a C-terminal domain, and a C-terminal tail; or (i) is the complement of the nucleotide sequence of (h). Preferably, for the BUB1 kinase, this domain is selected from the group N-terminal domain, catalytic domain, C-terminal catalytic domain, and C-terminal tail, and for NEK4a kinase, NEK4b kinase, NEK5 kinase or NEK6 kinase this domain is selected from the group N-terminal domain, catalytic domain, C-terminal catalytic domain, and C-terminal domain.

The term "complement" refers to two nucleotides that can form multiple favorable interactions with one another. For example, adenine is complementary to thymine as they can form two hydrogen bonds. Similarly, guanine and cytosine are complementary since they can form three hydrogen bonds. A nucleotide sequence is the complement of another nucleotide sequence if all of the nucleotides of the first sequence are complementary to all of the nucleotides of the second sequence.

The term "domain" refers to a region of a polypeptide which contains a particular function. For instance, N-terminal or C-terminal domains of signal transduction proteins can serve functions including, but not limited to, binding molecules that localize the signal

transduction molecule to different regions of the cell or binding other signaling molecules directly responsible for propagating a particular cellular signal. Some domains can be expressed separately from the rest of the protein and function by themselves, while others must remain part of the intact protein to retain function. The latter are termed functional regions of proteins and also relate to domains.

The term "N-terminal domain" refers to the extracatalytic region located between the initiator methionine and the catalytic domain of the protein kinase. The N-terminal domain can be identified following a Smith-Waterman alignment of the protein sequence against the non-redundant protein database to define the N-terminal boundary of the catalytic domain. Depending on its length, the N-terminal domain may or may not play a regulatory role in kinase function. An example of a protein kinase whose N-terminal domain has been shown to play a regulatory role is PAK65, which contains a CRIB motif used for Cdc42 and rac binding (Burbelo, P.D. *et al.* (1995) J. Biol. Chem. 270, 29071-29074). The N-terminal domain spans amino acid residues 1-785 of the sequence set forth in SEQ ID NO:2; 1-31 of the sequence set forth in SEQ ID NO:4; 1-22 of the sequence set forth in SEQ ID NO:6; and 1-42 of the sequence set forth in SEQ ID NO:8.

The term "catalytic domain" refers to a region of the protein kinase that is typically 25-300 amino acids long and is responsible for carrying out the phosphate transfer reaction from a high-energy phosphate donor molecule such as ATP or GTP to itself (autophosphorylation) or to other proteins (exogenous phosphorylation). The catalytic domain of protein kinases is made up of 12 subdomains that contain highly conserved amino acid residues, and are responsible for proper polypeptide folding and for catalysis. The catalytic domain can be identified following a Smith-Waterman alignment of the protein sequence against the non-redundant protein database. The catalytic domain spans amino acid residues 32-283 of the sequence set forth in SEQ ID NO:4; and 23-302 of the sequence set forth in SEQ ID NO:6; and 43-294 of the sequence set forth in SEQ ID NO:8; and 1-259 of the sequence set forth in SEQ ID NO:10. The C-terminal catalytic domain spans amino acid residues 786-1028 of the sequence set forth in SEQ ID NO:2.

The term "catalytic activity", as used herein, defines the rate at which a kinase catalytic domain phosphorylates a substrate. Catalytic activity can be measured, for example, by determining the amount of a substrate converted to a phosphorylated product as a function of time. Catalytic activity can be measured by methods of the invention by holding time constant and determining the concentration of a phosphorylated substrate

after a fixed period of time. Phosphorylation of a substrate occurs at the active-site of a protein kinase. The active-site is normally a cavity in which the substrate binds to the protein kinase and is phosphorylated.

The term "substrate" as used herein refers to a molecule phosphorylated by a kinase of the invention. Kinases phosphorylate substrates on serine/threonine or tyrosine amino acids. The molecule may be another protein or a polypeptide.

The term "C-terminal domain" refers to the region located between the catalytic domain or the last (located closest to the C-terminus) functional domain and the carboxy-terminal amino acid residue of the protein kinase. By "functional" domain is meant any region of the polypeptide that may play a regulatory or catalytic role as predicted from amino acid sequence homology to other proteins or by the presence of amino acid sequences that may give rise to specific structural conformations (*e.g.* N-terminal domain). The C-terminal domain can be identified by using a Smith-Waterman alignment of the protein sequence against the non-redundant protein database to define the C-terminal boundary of the catalytic domain or of any functional C-terminal extracatalytic domain. Depending on its length and amino acid composition, the C-terminal domain may or may not play a regulatory role in kinase function. An example of a protein kinase whose C-terminal domain may play a regulatory role is PAK3 which contains a heterotrimeric G_b subunit-binding site near its C-terminus (Leeuw, T. *et al.* (1998) Nature, 391, 191-195). For the some of the kinases of the instant invention, the C-terminal domain also comprises the catalytic domain (above). The C-terminal catalytic domain spans amino acid residues 786-1028 of the sequence set forth in SEQ ID NO:2; 284-302 of the sequence set forth in SEQ ID NO:4; 295-313 of the sequence set forth in SEQ ID NO:8; and 260-268 of the sequence set forth in SEQ ID NO:10.

The term "C-terminal tail" as used herein, refers to a C-terminal domain of a protein kinase, that by homology extends or protrudes past the C-terminal amino acid of its closest homolog. C-terminal tails can be identified by using a Smith-Waterman sequence alignment of the protein sequence against the non-redundant protein database, or by means of a multiple sequence alignment of homologous sequences using the DNAStar program Megalign. Depending on its length, a C-terminal tail may or may not play a regulatory role in kinase function. The C-terminal tail spans amino acid residues 1029-1085 of the sequence set forth in SEQ ID NO:2.

The term "signal transduction pathway" refers to the molecules that propagate an extracellular signal through the cell membrane to become an intracellular signal. This signal can then stimulate a cellular response. The polypeptide molecules involved in signal transduction processes are typically receptor and non-receptor protein tyrosine kinases, receptor and non-receptor protein phosphatases, SRC homology 2 and 3 domains, phosphotyrosine binding proteins (SRC homology 2 (SH2) and phosphotyrosine binding (PTB and PH) domain containing proteins), proline-rich binding proteins (SH3 domain containing proteins), nucleotide exchange factors, and transcription factors.

Various low or high stringency hybridization conditions may be used depending upon the specificity and selectivity desired. These conditions are well-known to those skilled in the art. Under stringent hybridization conditions only highly complementary nucleic acid sequences hybridize. Preferably, such conditions prevent hybridization of nucleic acids having more than 1 or 2 mismatches out of 20 contiguous nucleotides, more preferably, such conditions prevent hybridization of nucleic acids having more than 1 or 2 mismatches out of 50 contiguous nucleotides, most preferably, such conditions prevent hybridization of nucleic acids having more than 1 or 2 mismatches out of 100 contiguous nucleotides. In some instances, the conditions may prevent hybridization of nucleic acids having more than 5 mismatches in the full-length sequence.

By stringent hybridization assay conditions is meant hybridization assay conditions at least as stringent as the following: hybridization in 50% formamide, 5X SSC, 50 mM NaH_2PO_4, pH 6.8, 0.5% SDS, 0.1 mg/mL sonicated salmon sperm DNA, and 5X Denhardt's solution at 42 °C overnight; washing with 2X SSC, 0.1% SDS at 45 °C; and washing with 0.2X SSC, 0.1% SDS at 45 °C. Under some of the most stringent hybridization assay conditions, the second wash can be done with 0.1X SSC at a temperature up to 70 °C (Berger *et al.* (1987) Guide to Molecular Cloning Techniques pg 421, hereby incorporated by reference herein in its entirety including any figures, tables, or drawings.). However, other applications may require the use of conditions falling between these sets of conditions. Methods of determining the conditions required to achieve desired hybridizations are well-known to those with ordinary skill in the art, and are based on several factors, including but not limited to, the sequences to be hybridized and the samples to be tested.

In other preferred embodiments, the invention features isolated, enriched, or purified nucleic acid molecules encoding kinase polypeptides, further comprising a vector

or promoter effective to initiate transcription in a host cell. The invention also features recombinant nucleic acid, preferably in a cell or an organism. The recombinant nucleic acid may contain a sequence set forth in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, or SEQ ID NO:9, or a functional derivative thereof and a vector or a promoter effective to initiate transcription in a host cell. The recombinant nucleic acid can alternatively contain a transcriptional initiation region functional in a cell, a sequence complementary to an RNA sequence encoding a kinase polypeptide and a transcriptional termination region functional in a cell. Specific vectors and host cell combinations are discussed herein.

The term "vector" relates to a single or double-stranded circular nucleic acid molecule that can be transfected into cells and replicated within or independently of a cell genome. A circular double-stranded nucleic acid molecule can be cut and thereby linearized upon treatment with restriction enzymes. An assortment of nucleic acid vectors, restriction enzymes, and the knowledge of the nucleotide sequences cut by restriction enzymes are readily available to those skilled in the art. A nucleic acid molecule encoding a kinase can be inserted into a vector by cutting the vector with restriction enzymes and ligating the two pieces together.

The term "transfecting" defines a number of methods to insert a nucleic acid vector or other nucleic acid molecules into a cellular organism. These methods involve a variety of techniques, such as treating the cells with high concentrations of salt, an electric field, detergent, or DMSO to render the outer membrane or wall of the cells permeable to nucleic acid molecules of interest or use of various viral transduction strategies.

The term "promoter" as used herein, refers to nucleic acid sequence needed for gene sequence expression. Promoter regions vary from organism to organism, but are well known to persons skilled in the art for different organisms. For example, in prokaryotes, the promoter region contains both the promoter (which directs the initiation of RNA transcription) as well as the DNA sequences which, when transcribed into RNA, will signal synthesis initiation. Such regions will normally include those 5'-non-coding sequences involved with initiation of transcription and translation, such as the TATA box, capping sequence, CAAT sequence, and the like.

In preferred embodiments, the isolated nucleic acid comprises, consists essentially of, or consists of a nucleic acid sequence set forth in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, or SEQ ID NO:9 which encodes the amino acid sequence of SEQ

ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, a functional derivative thereof, or at least 35, 40, 45, 50, 60, 75, 100, 200, or 300 contiguous amino acids of SEQ ID NO:2, or at least 35, 40, 45, 50, 60, 75, 100, 200, or 300 contiguous amino acids of SEQ ID NO:4, or at least 10, 15, 20, 30, 40, 45, 50, 60, 75, 100, 200, or 300 contiguous amino acids of SEQ ID NO:6, or at least 35, 40, 45, 50, 60, 75, 100, 200, or 300 contiguous amino acids of SEQ ID NO:8, or at least 35, 40, 45, 50, 60, 75, 100, 200, or 300 contiguous amino acids of SEQ ID NO:10. A BUB1 kinase polypeptide comprises, consists essentially of, or consists of at least 35, 40, 45, 50, 60, 75, 100, 200, or 300 contiguous amino acids of SEQ ID NO:2. A NEK kinase polypeptide comprises, consists essentially of, or consists of at least 35, 40, 45, 50, 60, 75, 100, 200, or 300 contiguous amino acids of SEQ ID NO:4, at least 10, 15, 20, 30, 40, 45, 50, 60, 75, 100, 200, or 300 contiguous amino acids of SEQ ID NO:6, of at least 35, 40, 45, 50, 60, 75, 100, 200, or 300 contiguous amino acids of SEQ ID NO:8, or of at least 35, 40, 45, 50, 60, 75, 100, 200, or 300 contiguous amino acids of SEQ ID NO:10. The nucleic acid may be isolated from a natural source by cDNA cloning or by subtractive hybridization. The natural source may be mammalian, preferably human, blood, semen, or tissue, and the nucleic acid may be synthesized by the triester method or by using an automated DNA synthesizer.

The term "mammal" refers preferably to such organisms as mice, rats, rabbits, guinea pigs, sheep, and goats, more preferably to cats, dogs, monkeys, and apes, and most preferably to humans.

In yet other preferred embodiments, the nucleic acid is a conserved or unique region, for example those useful for: the design of hybridization probes to facilitate identification and cloning of additional polypeptides, the design of PCR probes to facilitate cloning of additional polypeptides, obtaining antibodies to polypeptide regions, and designing antisense oligonucleotides.

By "conserved nucleic acid regions", are meant regions present on two or more nucleic acids encoding a kinase polypeptide, to which a particular nucleic acid sequence can hybridize under lower stringency conditions. Examples of lower stringency conditions suitable for screening for nucleic acid encoding kinase polypeptides are provided in Abe, *et al.* (J. Biol. Chem. 19:13361-13368, 1992), hereby incorporated by reference herein in its entirety, including any drawings, figures, or tables. Preferably,

conserved regions differ by no more than 5 out of 20 nucleotides, even more preferably 2 out of 20 nucleotides or most preferably 1 out of 20 nucleotides.

By "unique nucleic acid region" is meant a sequence present in a nucleic acid coding for a kinase polypeptide that is not present in a sequence coding for any other naturally occurring polypeptide. Such regions preferably encode 32 (preferably 40, more preferably 45, most preferably 55) or more contiguous amino acids set forth in the full-length amino acid sequence of SEQ ID NO:2, 33 (preferably 40, more preferably 45, most preferably 55) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:4, or 8 (preferably 15, more preferably 20, most preferably 30) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:6, 33 (preferably 40, more preferably 45, most preferably 55) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:8, 33 (preferably 40, more preferably 45, most preferably 55) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:10. In particular, a unique nucleic acid region is preferably of mammalian origin.

A second aspect of the invention features a nucleic acid probe for the detection of nucleic acid encoding a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2 in a sample or a NEK kinase polypeptide in a sample. In preferred embodiments, the nucleic acid probe encodes a kinase polypeptide that is a fragment of the protein encoded by the amino acid sequence set forth in SEQ ID NO:4 or SEQ ID NO:6. Preferably, the NEK kinase polypeptide is selected from the group consisting of NEK4, NEK5, and NEK6; preferably NEK4 includes NEK4a having the amino acid sequence set forth in SEQ ID NO:8 and NEK4b having the amino acid sequence set forth in SEQ ID NO:10; more preferably the kinase polypeptide is selected from the group consisting of NEK5 having the amino acid sequence set forth in SEQ ID NO:4 and NEK6; and most preferably the kinase polypeptide is NEK 6 having the amino acid sequence set forth in SEQ ID NO:6. The nucleic acid probe contains a nucleotide base sequence that will hybridize to the sequence set forth in SEQ ID NO: 1, SEQ ID NO:3 or SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, or a functional derivative thereof.

In preferred embodiments, the nucleic acid probe hybridizes to nucleic acid encoding at least 12, 32, 75, 90, 105, 120, 150, 200, 250, 300 or 350 contiguous amino acids of the full-length sequence set forth in SEQ ID NO:2, at least 35, 75, 90, 105, 120, 150, 200, 250, 300 or 350 contiguous amino acids of the full-length sequence set forth in

SEQ ID NO:4, at least 12, 32, 75, 90, 105, 120, 150, 200, 250, 300 or 350 contiguous amino acids of the full-length sequence set forth in SEQ ID NO:6, at least 35, 75, 90, 105, 120, 150, 200, 250, 300 or 350 contiguous amino acids of the full-length sequence set forth in SEQ ID NO:8, at least 35, 75, 90, 105, 120, 150, 200, 250, 300 or 350 contiguous amino acids of the full-length sequence set forth in SEQ ID NO:10, or a functional derivative thereof.

Methods for using the probes include detecting the presence or amount of kinase RNA in a sample by contacting the sample with a nucleic acid probe under conditions such that hybridization occurs and detecting the presence or amount of the probe bound to kinase RNA. The nucleic acid duplex formed between the probe and a nucleic acid sequence coding for a kinase polypeptide may be used in the identification of the sequence of the nucleic acid detected (Nelson *et al.*, in Nonisotopic DNA Probe Techniques, Academic Press, San Diego, Kricka, ed., p. 275, 1992, hereby incorporated by reference herein in its entirety, including any drawings, figures, or tables). Kits for performing such methods may be constructed to include a container means having disposed therein a nucleic acid probe.

In a third aspect, the invention describes a recombinant cell or tissue comprising a nucleic acid molecule encoding a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2 or encoding a NEK kinase polypeptide. Preferably, the NEK kinase polypeptide is selected from the group consisting of NEK4, NEK5, and NEK6; preferably NEK4 includes NEK4a having the amino acid sequence set forth in SEQ ID NO:8, and NEK4b having the amino acid sequence set forth in SEQ ID NO:10; more preferably the kinase polypeptide is selected from the group consisting of NEK5 having the amino acid sequence set forth in SEQ ID NO:4 and NEK6 having the amino acid sequence set forth in SEQ ID NO:6; and most preferably the kinase polypeptide is NEK 6. In such cells, the nucleic acid may be under the control of the genomic regulatory elements, or may be under the control of exogenous regulatory elements including an exogenous promoter. By "exogenous" it is meant a promoter that is not normally coupled *in vivo* transcriptionally to the coding sequence for the kinase polypeptides.

The polypeptide is preferably a fragment of the protein encoded by the full-length amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10. By "fragment," is meant an amino acid sequence present in a kinase polypeptide. Preferably, such a sequence comprises at least 32, 45, 50, 60, 100,

200, or 300 contiguous amino acids of the full-length sequence set forth in SEQ ID NO:2, comprises at least 35, 45, 50, 60, 100, 200, or 300 contiguous amino acids of the sequence set forth in SEQ ID NO:4, comprises at least 10, 15, 20, 30, 45, 50, 60, 100, 200, or 300 contiguous amino acids of the sequence set forth in SEQ ID NO:6, comprises at least 35, 45, 50, 60, 100, 200, or 300 contiguous amino acids of the sequence set forth in SEQ ID NO:8, or comprises at least 35, 45, 50, 60, 100, 200, or 300 contiguous amino acids of the sequence set forth in SEQ ID NO:10.

In a fourth aspect, the invention features an isolated, enriched, or purified BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2 or an isolated, enriched, or purified NEK kinase polypeptide. Preferably, the NEK kinase polypeptide is selected from the group consisting of NEK4, NEK5, and NEK6; preferably NEK4 includes NEK4a having the amino acid sequence set forth in SEQ ID NO:8 and NEK4b having the amino acid sequence set forth in SEQ ID NO:10; more preferably the kinase polypeptide is selected from the group consisting of NEK5 having the amino acid sequence set forth in SEQ ID NO:4 and NEK6 having the amino acid sequence set forth in SEQ ID NO:6; and most preferably the kinase polypeptide is NEK 6.

By "isolated" in reference to a polypeptide is meant a polymer of 6 (preferably 12, more preferably 18, most preferably 25, 32, 40, or 50) or more amino acids conjugated to each other, including polypeptides that are isolated from a natural source or that are synthesized. In certain aspects longer polypeptides are preferred, such as those with 100, 200, 300, 400, or more contiguous amino acids of the full-length sequence set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10.

The isolated polypeptides of the present invention are unique in the sense that they are not found in a pure or separated state in nature. Use of the term "isolated" indicates that a naturally occurring sequence has been removed from its normal cellular environment. Thus, the sequence may be in a cell-free solution or placed in a different cellular environment. The term does not imply that the sequence is the only amino acid chain present, but that it is essentially free (about 90 - 95% pure at least) of non-amino acid-based material naturally associated with it.

By the use of the term "enriched" in reference to a polypeptide is meant that the specific amino acid sequence constitutes a significantly higher fraction (2- to 5-fold) of the total amino acid sequences present in the cells or solution of interest than in normal or diseased cells or in the cells from which the sequence was taken. This could be caused by

a person by preferential reduction in the amount of other amino acid sequences present, or by a preferential increase in the amount of the specific amino acid sequence of interest, or by a combination of the two. However, it should be noted that enriched does not imply that there are no other amino acid sequences present, just that the relative amount of the sequence of interest has been significantly increased. The term significant here is used to indicate that the level of increase is useful to the person making such an increase, and generally means an increase relative to other amino acid sequences of about at least 2-fold, more preferably at least 5- to 10-fold or even more. The term also does not imply that there is no amino acid sequence from other sources. The other source of amino acid sequences may, for example, comprise amino acid sequence encoded by a yeast or bacterial genome, or a cloning vector such as pUC19. The term is meant to cover only those situations in which man has intervened to increase the proportion of the desired amino acid sequence.

It is also advantageous for some purposes that an amino acid sequence be in purified form. The term "purified" in reference to a polypeptide does not require absolute purity (such as a homogeneous preparation); instead, it represents an indication that the sequence is relatively purer than in the natural environment. Compared to the natural level this level should be at least 2-to 5-fold greater (*e.g.*, in terms of mg/mL). Purification of at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated. The substance is preferably free of contamination at a functionally significant level, for example 90%, 95%, or 99% pure.

In preferred embodiments, the kinase polypeptide is a fragment of the protein encoded by the full-length amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10. Preferably, the kinase polypeptide contains at least 32, 45, 50, 60, 100, 200, or 300 contiguous amino acids of the full-length sequence set forth in SEQ ID NO:2, at least 35, 45, 50, 60, 100, 200, or 300 contiguous amino acids of the sequence set forth in SEQ ID NO:4, at least 10, 15, 25, 35, 45, 50, 60, 100, 200, or 300 contiguous amino acids of the sequence set forth in SEQ ID NO:6, at least 35, 45, 50, 60, 100, 200, or 300 contiguous amino acids of the sequence set forth in SEQ ID NO:8, or at least 35, 45, 50, 60, 100, 200, or 300 contiguous amino acids of the sequence set forth in SEQ ID NO:10, or a functional derivative thereof.

In preferred embodiments, the kinase polypeptide comprises an amino acid sequence having (a) the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4,

SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10; (b) the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6, except that it lacks one or more of the following segments of amino acid residues: 1-785, 786-1028, or 1029-1085 of SEQ ID NO:2, 1-31, 32-283, or 284-302 of SEQ ID NO:4, or 1-22 or 23-302 of SEQ ID NO:6; (c) the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6 from amino acid residues 1-785, 786-1028, or 1029-1085 of SEQ ID NO:2, 1-31, 32-283, or 284-302 of SEQ ID NO:4, or 1-22 or 23-302 of SEQ ID NO:6; or (d) the amino acid sequence set forth in SEQ ID NO:2, except that it lacks one or more of the domains selected from the group consisting of a C-terminal catalytic domain, an N-terminal domain, and a C-terminal tail, or the amino acid sequence set forth in SEQ ID NO:4 or SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10, except that it lacks one or more, but not all, of the domains selected from the group consisting of an N-terminal domain, a catalytic domain a C-terminal catalytic domain, and a C-terminal domain.

The polypeptide can be isolated from a natural source by methods well-known in the art. The natural source may be mammalian, preferably human, blood, semen, or tissue, and the polypeptide may be synthesized using an automated polypeptide synthesizer.

In some embodiments the invention includes a recombinant kinase polypeptide selected from the group consisting of BUB1, NEK4a, NEK4b, NEK5 and NEK6 kinase polypeptides. By "recombinant kinase polypeptide" is meant a polypeptide produced by recombinant DNA techniques such that it is distinct from a naturally occurring polypeptide either in its location (*e.g.*, present in a different cell or tissue than found in nature), purity or structure. Generally, such a recombinant polypeptide will be present in a cell in an amount different from that normally observed in nature.

In a fifth aspect, the invention features an antibody (*e.g.*, a monoclonal or polyclonal antibody) having specific binding affinity to a kinase polypeptide or a kinase polypeptide domain or fragment where the polypeptide is selected from the group consisting of a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2, a NEK4a kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:8, a NEK4b kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:10, a NEK5 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:4, and a NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6. By "specific binding affinity" is meant that the antibody binds to the target kinase polypeptide with greater affinity than it binds to other polypeptides under specified

conditions. Antibodies or antibody fragments are polypeptides that contain regions that can bind other polypeptides. The term "specific binding affinity" describes an antibody that binds to a kinase polypeptide with greater affinity than it binds to other polypeptides under specified conditions.. Antibodies can be used to identify an endogenous source of BUB1 or NEK kinase polypeptides, to monitor cell cycle regulation, and for immuno-localization of NEK kinase polypeptides to the centrosomes.

The term "polyclonal" refers to antibodies that are heterogenous populations of antibody molecules derived from the sera of animals immunized with an antigen or an antigenic functional derivative thereof. For the production of polyclonal antibodies, various host animals may be immunized by injection with the antigen. Various adjuvants may be used to increase the immunological response, depending on the host species.

"Monoclonal antibodies" are substantially homogenous populations of antibodies to a particular antigen. They may be obtained by any technique which provides for the production of antibody molecules by continuous cell lines in culture. Monoclonal antibodies may be obtained by methods known to those skilled in the art (Kohler *et al.*, Nature 256:495-497, 1975, and U.S. Patent No. 4,376,110, both of which are hereby incorporated by reference herein in their entirety including any figures, tables, or drawings).

The term "antibody fragment" refers to a portion of an antibody, often the hypervariable region and portions of the surrounding heavy and light chains, that displays specific binding affinity for a particular molecule. A hypervariable region is a portion of an antibody that physically binds to the polypeptide target.

Antibodies or antibody fragments having specific binding affinity to a kinase polypeptide of the invention may be used in methods for detecting the presence and/or amount of kinase polypeptide in a sample by probing the sample with the antibody under conditions suitable for kinase-antibody immunocomplex formation and detecting the presence and/or amount of the antibody conjugated to the kinase polypeptide. Diagnostic kits for performing such methods may be constructed to include antibodies or antibody fragments specific for the kinase as well as a conjugate of a binding partner of the antibodies or the antibodies themselves.

An antibody or antibody fragment with specific binding affinity to a kinase polypeptide of the invention can be isolated, enriched, or purified from a prokaryotic or eukaryotic organism. Routine methods known to those skilled in the art enable production

of antibodies or antibody fragments, in both prokaryotic and eukaryotic organisms. Purification, enrichment, and isolation of antibodies, which are polypeptide molecules, are described above.

Antibodies having specific binding affinity to a kinase polypeptide of the invention may be used in methods for detecting the presence and/or amount of kinase polypeptide in a sample by contacting the sample with the antibody under conditions such that an immunocomplex forms and detecting the presence and/or amount of the antibody conjugated to the kinase polypeptide. Diagnostic kits for performing such methods may be constructed to include a first container containing the antibody and a second container having a conjugate of a binding partner of the antibody and a label, such as, for example, a radioisotope. The diagnostic kit may also include notification of an FDA approved use and instructions therefor.

In a sixth aspect, the invention features a hybridoma which produces an antibody having specific binding affinity to a kinase polypeptide or a kinase polypeptide domain, where the polypeptide is selected from the group consisting of a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2, a NEK4a kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:8, a NEK4b kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:10, a NEK5 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:4, and a NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6. By "hybridoma" is meant an immortalized cell line that is capable of secreting an antibody, for example an antibody to a kinase of the invention. In preferred embodiments, the antibody to the kinase comprises a sequence of amino acids that is able to specifically bind a kinase polypeptide of the invention.

In a seventh aspect, the invention features a kinase polypeptide binding agent able to bind to a kinase polypeptide selected from the group consisting of a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2, a NEK4a kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:8, a NEK4b kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:10, a NEK5 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:4, and a NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6. The binding agent is preferably a purified antibody that recognizes an epitope present on a kinase polypeptide of the invention. Other binding agents include molecules that bind to

kinase polypeptides and analogous molecules that bind to a kinase polypeptide. Such binding agents may be identified by using assays that measure kinase binding partner activity, such as those that measure PDGFR activity.

The invention also features a method for screening for human cells containing a kinase polypeptide of the invention or an equivalent sequence. The method involves identifying the novel polypeptide in human cells using techniques that are routine and standard in the art, such as those described herein for identifying the kinases of the invention (*e.g.*, cloning, Southern or Northern blot analysis, in situ hybridization, PCR amplification, etc.).

In an eighth aspect, the invention features methods for identifying a substance that modulates kinase activity comprising the steps of: (a) contacting a kinase polypeptide selected from the group consisting of a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2, a NEK4a kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:8, a NEK4b kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:10, a NEK5 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:4, and a NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6 with a test substance; (b) measuring the activity of said polypeptide; and (c) determining whether said substance modulates the activity of said polypeptide.

The term "modulates" refers to the ability of a compound to alter the function of a kinase of the invention. A modulator preferably activates or inhibits the activity of a kinase of the invention depending on the concentration of the compound exposed to the kinase.

The term "activates" refers to increasing the cellular activity of the kinase. The term inhibit refers to decreasing the cellular activity of the kinase. Kinase activity is preferably the interaction with a natural binding partner.

The term "modulates" also refers to altering the function of kinases of the invention by increasing or decreasing the probability that a complex forms between the kinase and a natural binding partner. A modulator preferably increases the probability that such a complex forms between the kinase and the natural binding partner, more preferably increases or decreases the probability that a complex forms between the kinase and the natural binding partner depending on the concentration of the compound exposed to the

kinase, and most preferably decreases the probability that a complex forms between the kinase and the natural binding partner.

The term "complex" refers to an assembly of at least two molecules bound to one another. Signal transduction complexes often contain at least two protein molecules bound to one another. For instance, a protein tyrosine receptor protein kinase, GRB2, SOS, RAF, and RAS assemble to form a signal transduction complex in response to a mitogenic ligand.

The term "natural binding partner" refers to polypeptides, lipids, small molecules, or nucleic acids that bind to kinases in cells. A change in the interaction between a kinase and a natural binding partner can manifest itself as an increased or decreased probability that the interaction forms, or an increased or decreased concentration of kinase/natural binding partner complex.

The term "contacting" as used herein refers to mixing a solution comprising the test compound with a liquid medium bathing the cells of the methods. The solution comprising the compound may also comprise another component, such as dimethyl sulfoxide (DMSO), which facilitates the uptake of the test compound or compounds into the cells of the methods. The solution comprising the test compound may be added to the medium bathing the cells by utilizing a delivery apparatus, such as a pipette-based device or syringe-based device.

In a ninth aspect, the invention features methods for identifying a substance that modulates kinase activity in a cell comprising the steps of: (a) expressing a kinase polypeptide in a cell, wherein said polypeptide is selected from the group consisting of a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2, a NEK4a kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:8, a NEK4b kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:10, a NEK5 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:4, and a NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6; (b) adding a test substance to said cell; and (c) monitoring a change in cell phenotype or the interaction between said polypeptide and a natural binding partner.

The term "expressing" as used herein refers to the production of kinases of the invention from a nucleic acid vector containing kinase genes within a cell. The nucleic acid vector is transfected into cells using well known techniques in the art as described herein.

In a tenth aspect, the invention provides methods for treating a disease by administering to a patient in need of such treatment a substance that modulates the activity of a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2. Preferably, the disease is cancer, more preferably the cancer is selected from the group consisting of leukemia, cervical cancer, lymphoma, colon cancer, lung cancer, melanoma, ovarian cancer, CNS cancer, prostate cancer, kidney cancer, and breast cancer. Alternatively, the disease or disorder is CREST syndrome.

CREST syndrome is named for the primary symptoms of Calcinosis cutis, Raynaud's phenomenon, Esophageal involvement, Sclerodactyly, and Telangeictasia. Patients with CREST syndrome suffer from progressive systemic sclerosis, diffuse systemic scleroderma, primary pulmonary hypertension, idiopathic portal hypertension, as well as increased incidence of malignant cancers.

Substances useful for treatment of kinase-related disorders or diseases preferably show positive results in one or more *in vitro* assays for an activity corresponding to treatment of the disease or disorder in question (Examples of such assays are provided in the references in section VIII, below). Examples of substances that can be screened for favorable activity are provided and referenced in section VIII, below. The substances that modulate the activity of the kinases preferably include, but are not limited to, antisense oligonucleotides and inhibitors of protein kinases, as determined by methods and screens referenced in section VIII, below.

The term "preventing" refers to decreasing the probability that an organism contracts or develops an abnormal condition.

The term "treating" refers to having a therapeutic effect and at least partially alleviating or abrogating an abnormal condition in the organism.

The term "therapeutic effect" refers to the inhibition or activation factors causing or contributing to the abnormal condition. A therapeutic effect relieves to some extent one or more of the symptoms of the abnormal condition. In reference to the treatment of abnormal conditions, a therapeutic effect can refer to one or more of the following: (a) an increase in the proliferation, growth, and/or differentiation of cells; (b) inhibition (*i.e.*, slowing or stopping) of cell death; (c) inhibition of degeneration; (d) relieving to some extent one or more of the symptoms associated with the abnormal condition; and (e) enhancing the function of the affected population of cells. Compounds demonstrating efficacy against abnormal conditions can be identified as described herein.

The term "abnormal condition" with respect to the human BUB1 kinase refers to a function in the cells or tissues of an organism that deviates from their normal functions in that organism. An abnormal condition can relate to cell proliferation, cell differentiation, or cell survival. Abnormal conditions include leukemia, cervical, lymphoma, colon, lung, melanoma, ovarian, CNS, prostate, kidney, and breast cancer, as well as CREST syndrome.

Abnormal cell proliferative conditions include cancers such as fibrotic and mesangial disorders, abnormal angiogenesis and vasculogenesis, wound healing, psoriasis, diabetes mellitus, and inflammation, as well as the cancers of the invention.

Abnormal differentiation conditions include, but are not limited to neurodegenerative disorders, slow wound healing rates, and slow tissue grafting healing rates.

Abnormal cell survival conditions relate to conditions in which programmed cell death (apoptosis) pathways are activated or abrogated. A number of protein kinases are associated with the apoptosis pathways. Aberrations in the function of any one of the protein kinases could lead to cell immortality or premature cell death.

The term "aberration", in reference to the function of a kinase in a signal transduction process, refers to a kinase that is over- or under-expressed in an organism, mutated such that its catalytic activity is lower or higher than wild-type protein kinase activity, mutated such that it can no longer interact with a natural binding partner, is no longer modified by another protein kinase or protein phosphatase, or no longer interacts with a natural binding partner.

The term "administering" relates to a method of incorporating a compound into cells or tissues of an organism. The abnormal condition can be prevented or treated when the cells or tissues of the organism exist within the organism or outside of the organism. Cells existing outside the organism can be maintained or grown in cell culture dishes. For cells harbored within the organism, many techniques exist in the art to administer compounds, including (but not limited to) oral, parenteral, dermal, injection, and aerosol applications. For cells outside of the organism, multiple techniques exist in the art to administer the compounds, including (but not limited to) cell microinjection techniques, transformation techniques, and carrier techniques.

The abnormal condition can also be prevented or treated by administering a compound to a group of cells having an aberration in a signal transduction pathway to an

organism. The effect of administering a compound on organism function can then be monitored. The organism is preferably a mouse, rat, rabbit, guinea pig, or goat, more preferably a monkey or ape, and most preferably a human.

The eleventh aspect of the invention features methods for detection of nucleic acid encoding a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2 in a sample as a diagnostic tool for a disease or a disorder, wherein the method comprises the steps of: (a) contacting the sample with a nucleic acid probe which hybridizes under hybridization assay conditions to a nucleic acid target region of *bub1*, the probe comprising the nucleic acid sequence encoding the polypeptide, fragments thereof, and the complements of the sequences and fragments; and (b) detecting the presence or amount of the probe:target region hybrid as an indication of the disease or the disorder.

In preferred embodiments of the invention, the disease or disorder is cancer. More preferably the cancer is selected from the group consisting of leukemia, cervical cancer, lymphoma, colon cancer, lung cancer, melanoma cancer, ovarian cancer, CNS cancer, prostate cancer, kidney cancer, and breast cancer. Alternatively, the disease or disorder is CREST syndrome.

The nucleic acid "target region" is the full-length nucleotide base sequence set forth in SEQ ID NO:1, a functional derivative thereof, or a fragment thereof, to which the nucleic acid probe will specifically hybridize. Specific hybridization indicates that in the presence of other nucleic acids the probe only hybridizes detectably with the kinase of the invention's target region. Putative target regions can be identified by methods well known in the art consisting of alignment and comparison of the most closely related sequences in the database.

In preferred embodiments the nucleic acid probe hybridizes to a nucleic acid target region encoding at least 12, 32, 50, 75, 90, 105, 120, 150, 200, 250, 300 or 350 contiguous amino acids of the full-length sequence set forth in SEQ ID NO:2, or a functional derivative thereof. Hybridization conditions should be such that hybridization occurs only with the kinase genes in the presence of other nucleic acid molecules. Under stringent hybridization conditions only highly complementary nucleic acid sequences hybridize. Preferably, such conditions prevent hybridization of nucleic acids having 1 or 2 mismatches out of 20 contiguous nucleotides. Such conditions are defined *supra*.

The diseases for which detection of kinase genes in a sample could be diagnostic include diseases in which nucleic acids (DNA and/or RNA) encoding the BUB1 kinase are

amplified in comparison to normal cells. By "amplification" is meant increased numbers of kinase DNA or RNA in a cell compared with normal cells. In normal cells, kinases are typically found as single copy genes. In selected diseases, the chromosomal location of the kinase genes may be amplified, resulting in multiple copies of the gene, or amplification. Gene amplification can lead to amplification of kinase RNA, or kinase RNA can be amplified in the absence of kinase DNA amplification.

"Amplification" as it refers to RNA can be the detectable presence of kinase RNA in cells, since in some normal cells there is no basal expression of kinase RNA. In other normal cells, a basal level of expression of kinase exists, therefore in these cases amplification is the detection of at least 1- to 2-fold, and preferably more, kinase RNA, compared to the basal level.

The diseases that could be diagnosed by detection of nucleic acid encoding a BUB1 kinase polypeptide in a sample preferably include cancers and CREST syndrome. The test samples suitable for nucleic acid probing methods of the present invention include, for example, cells or nucleic acid extracts of cells, or biological fluids. The samples used in the above-described methods will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts to be assayed. Methods for preparing nucleic acid extracts of cells are well known in the art and can be readily adapted in order to obtain a sample which is compatible with the method utilized.

In a twelfth aspect, the invention provides methods for treating a disease by administering to a patient in need of such treatment a substance that modulates the activity of a NEK kinase polypeptide. Preferably, the kinase polypeptide is selected from the group consisting of NEK4, NEK5, and NEK6. Preferably, NEK4 includes NEK4a kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:8 and NEK4b kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:10; more preferably the kinase polypeptide is selected from the group consisting of NEK5 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:4 and NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6; and most preferably the kinase polypeptide is NEK 6.

Substances useful for treatment of kinase-related disorders or diseases preferably show positive results in one or more *in vitro* assays for an activity corresponding to treatment of the disease or disorder in question (Examples of such assays are provided in the references in section IX, and Examples 22-24, below). Examples of substances that

can be screened for favorable activity are provided and referenced in section IX, below. The substances that modulate the activity of the kinases preferably include, but are not limited to, antibodies, antisense oligonucleotides and inhibitors of protein kinases, as determined by methods and screens referenced in section IX and Examples 19, and 22-24, below. In preferred embodiments, the substance is an antisense oligonucleotide for the inhibition of the expression of NEK kinase polypeptide and fragments thereof. Preferably, the kinase polypeptide is selected from the group consisting of NEK4, NEK5, and NEK6. Preferably, NEK4 includes NEK4a and NEK4b, and the antisense oligonucleotides are synthesized as phosphorothioates.

The term "abnormal condition" with respect to a NEK kinase refers to a function in the cells or tissues of an organism that deviates from their normal functions in that organism. An abnormal condition can relate to cell proliferation, cell differentiation, or cell survival. Abnormal conditions include: cancer and neurodegenerative and immune disorders.

The thirteenth aspect of the invention features methods for detection of nucleic acid encoding a NEK kinase polypeptide in a sample as a diagnostic tool for a disease or a disorder, wherein said method comprises: (a) contacting the sample with a nucleic acid probe which hybridizes under hybridization assay conditions to a nucleic acid target region of the kinase polypeptide, the probe comprising the nucleic acid sequence encoding the kinase polypeptide, fragments thereof, and the complements of the sequences and fragments; and (b) detecting the presence or amount of the probe:target region hybrid as an indication of the disease or the disorder. Preferably, the kinase polypeptide is selected from the group consisting of NEK4, NEK5, and NEK6. Preferably, NEK4 includes NEK4a kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:8 and NEK4b kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:10; more preferably the kinase polypeptide is selected from the group consisting of NEK5 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:4 and NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6; and most preferably the kinase polypeptide is NEK 6.

In preferred embodiments of the invention, the disease or disorder is selected from the group consisting of cancer and neurodegenerative and immune disorders. More preferably the disease or disorder is cancer.

The nucleic acid "target region" is the full-length nucleotide base sequence set forth in SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, or SEQ ID NO:9, a functional derivative thereof, or a fragment thereof, to which the nucleic acid probe will specifically hybridize. Specific hybridization indicates that in the presence of other nucleic acids the probe only hybridizes detectably with the kinase of the invention's target region. Putative target regions can be identified by methods well known in the art consisting of alignment and comparison of the most closely related sequences in the database.

In preferred embodiments the nucleic acid probe hybridizes to a nucleic acid target region encoding at least 6, 12, 32, 50, 75, 90, 105, 120, 150, 200, 250, 300 or 350 contiguous amino acids of the sequence set forth in SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10, or a functional derivative thereof. Hybridization conditions should be such that hybridization occurs only with the kinase genes, even in the presence of other nucleic acid molecules. Under stringent hybridization conditions only highly complementary nucleic acid sequences hybridize. Preferably, such conditions prevent hybridization of nucleic acids having 1 or 2 mismatches out of 20 contiguous nucleotides. Such conditions are defined *supra*.

The diseases for which detection of kinase genes in a sample could be diagnostic include diseases in which NEK kinase polypeptide nucleic acid (DNA and/or RNA) is amplified in comparison to normal cells.

The diseases that could be diagnosed by detection of nucleic acid encoding a NEK kinase polypeptide in a sample preferably include cancer and neurodegenerative and immune disorders. The test samples suitable for nucleic acid probing methods of the present invention include, for example, cells or nucleic acid extracts of cells, or biological fluids. The samples used in the above-described methods will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts to be assayed. Methods for preparing nucleic acid extracts of cells are well known in the art and can be readily adapted in order to obtain a sample that is compatible with the method utilized.

The fourteenth aspect of the invention features antisense oligonucleotides for the inhibition of the expression of a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2 or a NEK kinase polypeptide. Preferably, the kinase polypeptide is selected from the group consisting of BUB1, NEK4, NEK5, and NEK6, and fragments thereof. Preferably, NEK4 includes NEK4a having the amino acid sequence set forth in SEQ ID NO:8 and NEK4b kinase polypeptide having the amino acid sequence set

forth in SEQ ID NO:10; more preferably the kinase polypeptide is selected from the group consisting of NEK5 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:4 and NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6; and most preferably the kinase polypeptide is NEK 6. In preferred embodiments, the oligonucleotide is the complement of a sequence encoding a fragment of the protein encoded by the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10. Preferably, the antisense oligonucleotides are synthesized as phosphorothioates.

Antisense oligonucleotides are preferably designed to specifically inhibit expression of a BUB1 or NEK kinases. Preferably this is accomplished by identifying sequences unique to kinase polypeptides of the invention (one or more). Preferably, these sequences are not present in other non-BUB1 or non-NEK kinase polypeptides present in a sample or in a patient. Methods to identify such sequences are well-known in the art and thus are only briefly described. Preferably, antisense oligonucleotides are 10 to 100 oligonucleotides, more preferably 15 to 30 nucleotides, and most preferably 18 to 25 nucleotides in length.

The antisense oligonucleotides of the invention are preferably used to inhibit BUB1 protein or NEK protein expression *in vivo* in normal and tumor cells. Antisense oligonucleotides can be used either singly or in combination. Preferably, expression of a BUB1 kinase or a NEK kinase is significantly reduced, and more preferably reduced to below the limit of detection. In preferred embodiments, treatment with the antisense oligonucleotides of the invention inhibits growth and/or induces apoptosis in cells. Antisense oligonucleotides can also be used to inhibit BUB1 or NEK protein expression in human tumor cell xenografts in nude mice. In preferred embodiments, antisense oligonucleotides are used as a treatment in various human diseases and disorders in which BUB1 kinase polypeptides or NEK kinase polypeptides are overexpressed or aberrantly expressed.

Additional antisense oligonucleotides and effective combinations can be identified by methods well known in the art. Briefly, cells or tissues overexpressing BUB1 or NEK kinase polypeptides can be contacted with antisense oligonucleotides, either singly or in combination, and the expression of *bub1* or *nek* RNA, and/or BUB1 kinase or NEK kinase polypeptides can be determined by methods described herein. Preferably, treatment with antisense oligonucleotides to BUB1 kinases or NEK kinases causes a decrease in the

expression of *bub1* or *nek* RNA and/or BUB1 or NEK polypeptides, more preferably expression is decreased significantly (1- to 2-fold), most preferably expression is decreased to an undetectable level.

The summary of the invention described above is not limiting and other features and advantages of the invention will be apparent from the following detailed description of the invention, and from the claims.

BRIEF DESCRIPTION OF THE FIGURES

Figures 1A-C show a human *bub1* nucleic acid sequence (SEQ ID NO:1).

Figure 2 shows a human BUB1 amino acid sequence (SEQ ID NO:2).

Figures 3A & B show a human BUB1 (SEQ ID NO:2), mouse BUB1 (SEQ ID NO:22), and *S. cerevisiae* BUB1 (SEQ ID NO:23) multiple sequence alignment.

Figures 4A, 4B, 4C, and 4D show the partial or full-length nucleic acid sequences of human NEK4a (SEQ ID NO:7), human NEK4b (SEQ ID NO:9), human NEK5 (SEQ ID NO:3), and murine NEK6 (SEQ ID NO:5) kinase polypeptides.

Figure 5 shows the partial or full-length amino acid sequences of human NEK4a (SEQ ID NO:8), human NEK4b (SEQ ID NO:10), human NEK5 (SEQ ID NO:4), and murine NEK6 (SEQ ID NO:6) kinase polypeptides.

Figure 6 shows a multiple sequence alignment of the amino acid sequences of NEK-subfamily kinases, including: human NEK4b (SEQ ID NO:10), human NEK5 (SEQ ID NO:4), mouse NEK6 (SEQ ID NO:6), *Emericella nidulans* NIMA (SEQ ID NO:11), human NEK2 (SEQ ID NO:12), and human GAK (SEQ ID NO:13) kinase polypeptides. The C-termini of NIMA, NEK2 and GAK are not shown in this figure, since NEK4, NEK5, and NEK6 lack long C-terminal regions.

Figure 7 shows a sequence alignment of the amino acid sequences of NEK4a (SEQ ID NO:8) and NEK4b (SEQ ID NO:10) kinase polypeptides.

DETAILED DESCRIPTION OF THE INVENTION

The Nucleic Acids of the Invention

A. Mammalian BUB1

The full length human *bub1* cDNA (SEQ ID NO:1) is 3408 bp long and consists of a 3255 bp ORF (45-3300) flanked by a 44 bp (1-44) 5' UTR and a 108 bp 3' UTR. A

potential polyadenylation signal begins at position 3404. The sequence flanking the first ATG conforms to the Kozak consensus (Kozak, M., Nucleic Acids Res. 15, 8125-8148 (1987)) for an initiating methionine, and is believed to be the translational start site for the human BUB1.

Two human *bub1* isoforms were originally identified as PCR fragments: SuSTK34a (R51-40-18, 238 bp), and SuSTK34b (R51-40-4, 246 bp). SuSTK34b is identical to SuSTK34a except for the presence of a 15 nucleotide deletion within kinase subdomain VII, just N-terminal to the typically conserved "DFG" motif in the activation loop. This isoform predicts the in frame deletion of the 5 amino acid sequence, "GLALI".

Eight human EST fragments are present in the dbEST database: R94348 spans human *bub1* cDNA from nucleotide 898-1259; the other seven span nucleotides 2049-3422 of human *bub1*.

B. Mammalian NEK4a

The full-length human NEK4a cDNA (SEQ ID NO:7) is 1,449 bp long and consists of a 1,107 bp open reading frame (ORF) flanked by a 121 bp 5' UTR and a 200 bp 3' UTR that is followed by a 21 nucleotide polyadenylated region. No polyadenylation signal (AATAAA) was found preceding the polyadenylated region. The sequence flanking the first ATG conforms to the Kozak consensus (Kozak, M. (1987) Nucleic Acids Res. 15:8125-8148) for an initiating methionine, and is believed to be the translational start site for NEK4a. A NEK4a isoform (NEK4b) was also identified whose sequence diverges from NEK4a beginning at position 1005 as a result of the substitution of the terminal 75 amino acids of NEK4a with a unique 19 amino acid C-terminal sequence.

Several EST fragments span the complete NEK4a sequence, with AA214523 at the 5' end and AA147800 at the 3' end.

C. Mammalian NEK4b

The partial human NEK4b cDNA (SEQ ID NO:9) is 1,346 bp long, and consists of a 804 bp ORF flanked by a 499 bp 3' UTR that is followed by a 41 nucleotide polyadenylated region. No polyadenylation signal (AATAAA) was found preceding the polyadenylated region. The first 750 nucleotides of NEK4b are identical to NEK4a, suggesting that these genes are alternatively spliced isoforms of a single NEK4 gene.

Several EST fragments span the NEK4b sequence, with N30324 at the 5' end and AA461376 at the 3' end.

D. Mammalian NEK5

The full-length human NEK5 cDNA (SEQ ID NO:3) is 3,854 bp long and consists of a 906 bp ORF flanked by a 31 bp 5' UTR and a 2,897 bp 3' UTR that is followed by a 20 nucleotide polyadenylated region. A polyadenylation signal (AATAAA) is found at positions 3817-3822. The sequence flanking the first ATG conforms to the Kozak consensus for an initiating methionine, and is believed to be the translational start site for NEK5.

At least one EST fragment, AA181165, spans the complete NEK5 sequence.

E. Mammalian NEK6

The partial murine NEK6 cDNA (SEQ ID NO:5) is 1,310 bp long and consists of a 907 bp ORF preceded by a 403 bp 5' UTR. The sequence flanking the first ATG only weakly conforms to the Kozak consensus for an initiating methionine, but is believed to be the translational start site for NEK5. There is an adenine at the -3 position prior to the start codon, but no pyrimidine at the +4 position. There are multiple in-frame and out-of-frame stop codons 5' of the putative start site, and no other possible start codons in any frame. The sequence preceding the start site is characteristic of a 5'UTR region, since it is GC-rich (57.6%).

One EST fragment, AA030068, represents the partial NEK6 sequence.

The Proteins of the Invention

A. Mammalian BUB1

The 3408 bp human *bub1* nucleotide sequence encodes a polypeptide of 1085 amino acids (SEQ ID NO:2). Analysis of the deduced amino acid sequence predicts BUB1 protein to be an intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain. BUB1 protein contains a large N terminal region (amino acid 1 to 785) followed by a C-terminal kinase domain (amino acid 786 to 1028) and a small C-terminal tail (amino acid 1029 to 1085) (FIG 3). The 785 amino acid N-terminal domain of human BUB1 shares 68% amino acid identity to murine BUB1. A 140 amino acid subdomain in this N-terminal region of human BUB1 shares 81%, 31%, and 29% amino acid sequence identity to murine BUB1 (AF002823) *Saccharomyces cerevisiae* BUB1 (L32027), and to the N-terminal domain of the non-kinase *S. cerevisiae* MAD3 protein (P47074), respectively. The 243 amino acid predicted catalytic domain of human BUB1 shares 83% and 36% amino acid identity to murine BUB1 and *S. cerevisiae* BUB1,

respectively. The 57 amino acid C-terminal tail of human BUB1 exhibits 72% and 24% amino acid identity to murine BUB1 and *S. cerevisiae* BUB1, respectively.

B. Mammalian NEK4a (Long C-terminal Isoform)

The 1,449 bp human NEK4a nucleotide sequence encodes a polypeptide of 369 amino acids (SEQ ID NO:8) with a predicted molecular mass of 41,825 daltons. Analysis of the deduced amino acid sequence predicts NEK4a to be an intracellular serine/threonine kinase, lacking both a signal sequence and a transmembrane domain. NEK4a contains a 42 amino acid N-terminal domain, a 252 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase, and a 75 amino acid C-terminal domain.

NEK4a is most closely related to human NEK5 (SEQ ID NO:4), a *C. elegans* kinase F19H6.1 (GB:Z68115), murine NEK1 (GB:S45828), human NEK3 (GB:Z29067), human NEK2 (GB:U11050), and *Emericella nidulans* NIMA (GB:M20249), sharing 82.6%, 78.8%, 41.2%, 40.2%, 44.2%, 42.6%, and 30.4% amino acid identity, respectively. NEK4a also shares 97.2% identity with human HPK1, a partial protein kinase disclosed in the International Publication No. WO98/11234.

NEK4a (369 amino acids) and NEK4b (268 amino acids) have identical amino acid (and nucleotide) sequence up to amino acids 304 and 249 in NEK4a and NEK4b, respectively, where these two proteins diverge, most likely as result of alternative splicing (FIG. 10).

The 42 amino acid N-terminal domain of human NEK4a does not reveal any significant homologies following a Smith-Waterman search of the non-redundant protein database. Human NEK4a lacks a glycine residue at position 2, and is therefore unlikely to undergo myristylation.

The 252 amino acid catalytic domain of human NEK4a is most closely related to human NEK5 (SEQ ID NO:4), a *C. elegans* kinase F19H6.1 (GB:Z68115), murine NEK1 (GB:S45828), human NEK3 (GB:Z29067), human NEK2 (GB:U11050), and *E. nidulans* NIMA (GB:M20249), sharing 88.5%, 78.8%, 41.6%, 40.6%, 44.2%, 42.6%, and 30.2% amino acid identity, respectively. NEK4a contains the potential "TPY" motif regulatory phosphorylation site in its activation loop. This "TPY" motif is also present in NEK5. NEK4a and NEK5 also contain the catalytic motif HPN(I/V) located C-terminal to the conserved catalytic lysine (position 102-105 in NEK4a, FIG.9) that is characteristic of the NIMA/NEK-subfamily of STK's. The kinase domain of NEK4a also shares 97.2%

identity with human HPK1, a partial protein kinase disclosed in the International Publication No. WO98/11234.

The 75 amino acid C-terminal domain of human NEK4a is most closely related to human NEK5 (SEQ ID NO:4), sharing 61.5% amino acid identity. The C-terminus of NEK4a does not display any detectable homology with human HPK1, a partial protein kinase disclosed in the International Publication No. WO98/11234.

In addition to the potential phosphorylation site in the "TPY" activation loop, the following additional potential phosphorylation sites are conserved in NEK4 and NEK5, as well as in their *C. elegans* orthologue, at the indicated positions in NEK4a (FIG. 9): protein kinase C [ST]-X-[RK] (PROSITE:PD0C00005) at 165; casein kinase II [ST-X(2)-[DE] (PROSITE:PD0C00006) at 111; tyrosine kinase [RK]-X-(2)-[DE]-X-(2/3)-Y (PROSITE:PD0C00007) at 224; and Cdc2 kinase [S/T]-P (Nigg, E. (1993) Trends in Cell Biol. 3:296-301) at 245. The 75 amino acid C-terminal region unique to NEK4a contains additional potential phosphorylation sites at 308 (Cdc2) and 344 (casein kinase II).

C. Mammalian NEK4b (Short C-terminal Isoform)

NEK4a (369 amino acids) and NEK4b (268 amino acids) have identical amino acid (and nucleotide) sequence up to amino acids 304 and 249 in NEK4a and NEK4b, respectively, where these two proteins diverge, most likely as result of alternative splicing (FIG. 4). This divergence occurs immediately following the predicted kinase domain, causing these two proteins to have divergent C-terminal tails.

The 19 amino acid C-terminal domain of human NEK4b is most related to human NEK5 (SEQ ID NO:4), sharing 61.5% amino acid identity. The C-terminal domain of NEK4b also shares 89.5% identity with human HPK1, a partial protein kinase disclosed in the International Publication No. WO98/11234.

D. Mammalian NEK5

The 3,854 bp human NEK5 nucleotide sequence encodes a polypeptide of 302 amino acids (SEQ ID NO:4) with a predicted molecular mass of 34,550 daltons. Analysis of the deduced amino acid sequence predicts NEK5 to be an intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain. NEK5 contains a 31 amino acid N-terminal domain, a 252 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase, and a 19 amino acid C-terminal domain.

NEK5 is most closely related to human NEK4a (SEQ ID NO:8), a *C. elegans* kinase F19H6.1 (GB:Z68115), murine NEK1 (GB:S45828), human NRK2 (GB:P51957),

human NEK3 (GB:Z299067), and *E. nidulans* NIMA (GB:M20249), sharing 82.6%, 79.8%, 38.3%, 40.2%, 44.0%, and 32.2% amino acid identity, respectively. NEK5 also shares 86.1% identity with human HPK1, a partial protein kinase disclosed in the International Publication No. WO98/11234.

The 31 amino acid N-terminal domain of human NEK5 is 34.6% identical to the N-terminus of a *C. elegans* kinase F19H6.1 (GB:Z68115). Human NEK5 lacks a glycine residue at position 2, and is therefore unlikely to undergo myristylation. A Smith-Waterman search of the nonredundant protein database does not reveal any significant homologies that might suggest a potential function for this domain.

The 252 amino acid catalytic domain of human NEK5 is most related to human NEK4a (SEQ ID NO:8), a *C. elegans* kinase F19H6.1 (GB:Z68115), murine NEK1 (GB:S45828), human NEK2 (GB:U11050), human NRK2 (GB:P51957), human NEK3 (GB:Z299067), and *E. nidulans* NIMA (GB:M20249), sharing 88.5%, 79.8%, 39.0%, 42.3%, 40.2%, 44.1%, and 32.2% amino acid identity, respectively. NEK5 contains the potential "TPY" motif regulatory phosphorylation site in its activation loop. This "TPY" motif is also present in NEK4a. NEK4a and NEK5 also share with NIMA, as well as with other NEK-subfamily kinases, the catalytic motif HPN(I/V) located C-terminal to the conserved catalytic lysine (position 102-105 in NEK4a, FIG. 9). The kinase domain of NEK5 shares 88.1% identity with human HPK1, a partial protein kinase disclosed in International Publication No. WO98/11234.

The 19 amino acid C-terminal domain of human NEK5 is most related to human NEK4a (SEQ ID NO:8), sharing 61.5% amino acid identity.

In addition to the potential phosphorylation site in the "TPY" activation loop, the following potential phosphorylation sites are conserved in NEK4 and NEK5, as well as in their *C. elegans* orthologue, at the indicated positions in NEK4a (FIG. 9): PKC at 165, casein kinase II at 111, tyrosine kinase at 224, and Cdc2 kinase at 245.

E. Mammalian NEK6

The 1,310 bp murine NEK6 partial cDNA nucleotide sequence encodes a polypeptide of 302 amino acids (SEQ ID NO:6), with a predicted molecular mass of 33,590 daltons. Analysis of the deduced amino acid sequence predicts NEK6 to be an intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain. NEK6 contains a 22 amino acid N-terminal domain, and a 280 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase. The NEK6

sequence is partial and, therefore, does not contain the C-terminal end of the catalytic domain nor the carboxyl-terminus encoded by the gene.

NEK6 is most closely related to a *Drosophila heteroneura* kinase (GB:AF052296), a *C. elegans* kinase (GB:Z46242), a *Saccharomyces cerevisiae* kinase (GB:YILO95W), human GAK (GB:D88435), human NEK1 (GB:S45828), human NEK3 (GB:Z299067), and human NRK (GB:L20321), sharing 67.8%, 43.7%, 37.0%, 39.5%, 30.5%, 33.9%, and 28.4% amino acid identity, respectively.

The 22 amino acid N-terminal domain of murine NEK6 lacked any significant homologies following a Smith-Waterman search of the non-redundant protein database.

The 281 amino acids of the catalytic domain of murine NEK6 are most related to a *D. heteroneura* kinase (GB:AF052296), a *C. elegans* kinase (GB:Z46242), a *S. cerevisiae* kinase (GB:YILO95W), human GAK (GB:D88435), human NEK1 (GB:S45828), human NEK3 (GB:Z299067), and human NRK (GB:L20321), sharing 67.8%, 43.7%, 37.0%, 39.5%, 30.5%, 33.9%, and 28.4% amino acid identity, respectively.

NEK6 shares with other NEK6-subfamily kinases three features that differentiate them from other NEK-subfamily members: (1) an insert between the catalytic glycine-rich loop and the catalytic lysine; (2) a three amino acid insert in the activation loop; and (3) the lack of a "TPY" motif regulatory phosphorylation motif in the activation loop.

NEK6 contains the following potential nucleotide sequence phosphorylation sites at the indicated positions: cAMP-dependent kinase [RK](2)-X-[ST] (PROSITE: PS00004) at 221; PKC at 9, 199, 226, 235, and 242; and casein kinase II at 7, 47, 115, 147, 241, and 278. Of these, the PKC site at 199 (two amino acids C-terminal to the "DFG" catalytic motif) is conserved across species (from plants to mammals) among NEK6-related kinases (murine and *Drosophila* NEK; *A. thaliana*, rat, and human GAK; and the *Saccharomyces* NEK-related kinases such as YILO95W). A significant, yet lower, degree of conservation is seen for the PKC site at position 235.

Biological Significance of hBUB1 Serine/Threonine Protein Kinase

Much attention has recently been paid to the relationship between cell cycle regulation and cancer development. Only a few genes involved in cell cycle regulation have been discovered in human cells. Examples include: p53, a tumor suppressor gene; ATM, a gene involved in an autoimmune disorder leading to a high incidence of cancer; and the SDI1 gene, which causes cell senescence. All three genes are cell cycle checkpoint genes which function in the G1/S transition.

A group of yeast cell cycle checkpoint genes was demonstrated in the G2-M transition and mitosis including BUB1, BUB2, BUB3, MAD1, MAD2, MAD3, and MPS1.

Little is known about mammalian proteins that control the G2-M transition or that monitor the spindle or centrosomes. However, evidence suggests that defects in the regulation of these processes could be important in tumorgenesis.

Human BUB1 is a novel gene encoding a serine/threonine kinase apparently functioning as a mitotic checkpoint. Human BUB1 is a centromeric protein which interacts with microtubules during the process of mitotic division.

Improper functioning of the mitotic spindle at metaphase can arrest cell cycle progression. For example, chromosome segregation at the metaphase-anaphase transition is prevented if one or more chromosomes are not yet congressed at the metaphase plate. Moreover, the initiation of a new cell cycle is prevented if mitosis was not completed in the previous cell cycle because of inhibition of microtubule assembly. Mitotic defects can cause genomic changes (aneuploid or tetraploid) and chromosome aberrations. Studying the biological function of the hBUB1 gene will improve the understanding of mitotic checkpoint control and cancer development.

Biological Significance of Novel NEK-Subfamily Kinases

A. NEK4 and NEK5

Human NEK4 and NEK5 are most closely related to the *C. elegans* kinase F19H6.1, and to four mammalian NEK-subfamily kinases: murine NEK1, and human NRK2, NEK2 and NEK3. They also demonstrate distant homology to the *E. nidulans* NIMA kinase.

Murine NEK1 (NIMA Related Kinase) is a 774 amino acid STK with a predicted molecular weight of 88.4 kDa. NEK1 has a 526 amino acid extracatalytic C-terminal

domain containing four putative nuclear-localization signals. The percent amino acid sequence identity between NEK1 and NIMA over the catalytic and extracatalytic regions is 42% and 14%, respectively.

NEK1 is expressed predominantly in male and female germ cells, suggesting a role for this kinase in meiosis. In addition, NEK1 is enriched in postmitotic neurons from peripheral ganglia, indicative of a role for this STK in differentiated cells.

NEK1 behaves as a dual-specificity kinase, phosphorylating serine/threonine and also tyrosine residues *in vitro*. Dual phosphorylation *in vivo* remains to be demonstrated (Arama, E. *et al.* (1998) Oncogene 16:1813-1823)

Human NRK2 (NIMA-related kinase 2) is an 841 amino acid STK with a predicted molecular weight of 94.6 kDa. NRK2 contains a 580 amino acid extracatalytic C-terminal domain whose sequence is unrelated to the C-termini found in NIMA and other NEK-subfamily kinases. The percent amino acid sequence identity between NRK2 and NIMA over the catalytic region is 39%.

NRK2 is expressed predominantly in liver, skeletal muscle, kidney, and pancreas. In contrast to NIMA, NRK2 transcript levels do not vary throughout the cell cycle. The NRK2 gene localizes to chromosome 3p21.1, a region that is frequently altered in renal carcinomas (Levedakou, E. *et al.* (1994) Oncogene 9:1977-1988).

Human NEK2 is a 445 amino acid STK with a predicted molecular weight of 51.7 kDa. The percent amino acid sequence identity between NEK2 and NIMA over the catalytic region and C-terminal extracatalytic regions is 47% and 19%, respectively.

NEK2 expression is highest in germ cells where it is found in meiotic pachytene-stage spermatocytes and oocytes. NEK2 levels increase at the G2/M transition of the meiotic cell cycle (Rhee, K and Wolgemuth, D. (1997) Development 124:2167-2177). NEK2 may also have a mitotic role, since the levels and activity of this kinase are maximal during the S and G2 phases of the cell cycle. A role in centrosome separation during mitosis has been postulated for NEK2 (Fry, A. *et al.* (1998) EMBO J. 17:470-481). The NEK2 gene is localized to human chromosome 14q12 (Schultz, S. *et al.* (1994) Cell Growth and Diff. 5:625-635).

Human NEK3 is a 459 amino acid STK with a predicted molecular weight of 52.3 kDa. The database entry for NEK3 (GB:Z29067) is missing the N-terminal catalytic-region and start methionine of this protein. The percent amino acid sequence identity

between NEK3 and NIMA over the catalytic region and C-terminal extracatalytic regions is 43% and 19%, respectively (Schultz, S. *et al.* (1994) Cell Growth and Diff. 5:625-635).

NIMA, an STK found in the filamentous fungi, *Aspergillus* and *E. nidulans*, is a 699 amino acid protein with a predicted molecular weight of 78.9 kDa. Mitosis in *A. nidulans* requires both Cdc2 and NIMA.

NIMA expression and kinase activity are coordinately regulated during the cell cycle, being lowest in G1 and highest through the S and G2 phases. Full NIMA activation requires multiple phosphorylation events, some of which may be carried out by Cdc2. Exit from mitosis into G1 requires the inactivation of both NIMA and Cdc2, possibly through ubiquitin-mediated proteolysis. The C-terminus of NIMA displays consensus Cdc2 phosphorylation sites as well as PEST motifs that may mediate the function of this protein through phosphorylation and proteolysis.

Evidence for a NIMA-like pathway in mitosis in higher eukaryotes comes from the observation that expression of a dominant-negative form of NIMA in Hela cells causes an arrest in G2, and that overexpression of NIMA in *Aspergillus*, as well as in higher eukaryotes, triggers chromatin condensation (Fry, A. and Nigg, E.A. (1995) Current Biology, 5:1122-1125).

NEK4 and NEK5 contrast sharply with all other NEK-subfamily kinases in that they lack long extracatalytic C-termini. The function of these extracatalytic regions in NIMA and NEK2 is to regulate their catalytic activity and turnover in a cell cycle-dependent manner. NEK4 and NEK5 are predicted to have longer half-lives than other NEK-subfamily kinases because they lack motifs that target proteolysis such as those found in NIMA and NEK2.

The short C-terminal extracatalytic tail found in NEK4a could play a regulatory role in kinase activity and/or substrate recognition.

Modulation of NEK4 function may be carried out in part through alternative splicing of the C-terminus generating the NEK4a and NEK4b isoforms. NEK4 and NEK5 may also be modulated through phosphorylation, as these proteins have multiple conserved consensus sites for phosphorylation by various kinases.

NEK4 and NEK5 are predicted to be small, compact kinases akin to MAPK's and CDK's. MAPK's are regulated through phosphorylation of the "TXY" residues in the activation loop by dual-specificity MAPK kinases (MEK's). These phosphorylations reorient the activation loop, resulting in over a 1000-fold enzymatic activation. The

mechanism of this rearrangement involves the interaction of the 50 amino acid C-terminus of MAPK (α-helix L16) with the phosphorylated residues in the activation loop. The rearranged activation loop forms a tight pocket (P+1 loop) that defines the substrate specificity of MAPK's as proline-directed kinases capable of phosphorylating only those substrates whose phosphorylation serine/threonine residues are followed by a proline (Canagarajah B.J. *et al.* (1997) Cell, 90:859-869).

Three features of NEK4 and NEK5 suggest that they may behave like MAPK's: (1) their predicted small, compact nature; (2) the presence of the potential dual phosphorylation "TPY" site in the activation loop; and (3) the presence of short C-termini (19 amino acids in NEK4b and NEK5 and 75 amino acids in NEK4a) that could play a role similar to the 50 amino acid C-terminal α-helix L16 of MAPK in generating an active site.

NEK4 and NEK5, like MAPK, may also be subject to activation by MEK's through dual phosphorylation in their activation loop. In addition, NEK4 and NEK5, like MAPK's and CDK's, may be proline-directed kinases, a frequently observed specificity among kinases that phosphorylate transcription and cell cycle-regulatory molecules.

NEK4 and NEK5 define a new subclass of NEK-subfamily kinases that may potentially transduce signals originated at the level of activated membrane receptors to the nucleus where their targets could be transcriptional or cell-cycle modulator proteins.

B. NEK6

Murine NEK6 is most closely related to the *D. heteroneura* kinase AF052296, to the *S. cerevisiae* kinase (GB:YILO95W), and to the mammalian kinases GAK, NEK1, NEK3 and NRK2, with distal homology to *E. Nidulans* NIMA.

Human GAK is a 1,311 amino acid STK with a predicted molecular weight of 143.1 kDa. GAK contains a 998 amino acid extracatalytic C-terminal domain whose sequence is unrelated to the C-termini found in NIMA and in other NEK-subfamily kinases. The C-terminus of GAK contains a tensin/auxilin-like domain. GAK expression and kinase activity oscillate during the cell cycle with maximum levels found in G1. The GAK gene maps to human chromosome 4p16 (Kimura, S.H. *et al.* (1997) Genomics, 44:179-187).

Cyclin G is a member of the cyclin family of regulatory subunits of the cyclin-dependent family of STK's. Unlike cyclins A and B, which regulate the G2/M transition, cyclin G regulates the G1 to S phase transition of the cell cycle (Tamura, K. *et al.* 1993)

Oncogene, 8:2113-2118). Cyclin G is a transcriptional target of the p53 tumor suppressor protein. Its overexpression may override a p53-mediated cell cycle checkpoint resulting in a hypersensitivity to the cytotoxic effects of cancer chemotherapy (Smith, M.L. *et al.* (1997) Exp. Cell Res. 230:61-68).

Based on the sequence homology between NEK6 and GAK, NEK6 may also have a long C-terminal region, possibly containing a tensin/auxilin-like domain like GAK. At least three regions unique to the kinase domain of NEK6 are likely to exert a strong influence in its substrate specificity: (1) NEK6, and related kinases like human GAK, the *S. cerevisiae* kinases YNL020C, YIL095W ,and YBR059C, and the *C. elegans* ZA6242, have in common, relative to other STK's, the presence of a variable-length insert located between the glycine-rich loop (subdomain I) and the conserved nucleotide-binding lysine residue (subdomain II) (FIG. 3, only NEK6 and GAK shown); (2) NEK6 and its related kinases display a six amino acid insert between the "DFG" (subdomain VII) and "TPF" activation loop (subdomain VIII) catalytic motifs; and (3) the activation loop of NEK6 and its related kinases lack the "TPY" motif observed in NEK4 and NEK5 and in its place contain two contiguous uniformly conserved threonine residues.

NEK6 defines another subclass of NEK-subfamily kinases that are GAK-like, and consequently may be subject to modulation by regulatory subunits such as cyclins and by domains that may be found within its extracatalytic region.

Nucleic Acid Probes, Methods, and Kits for Detection of BUB1-Related and NEK Kinases

A nucleic acid probe of the present invention may be used to probe an appropriate chromosomal or cDNA library by usual hybridization methods to obtain other nucleic acid molecules of the present invention. A chromosomal DNA or cDNA library may be prepared from appropriate cells according to recognized methods in the art (*cf.* "Molecular Cloning: A Laboratory Manual", second edition, Cold Spring Harbor Laboratory, Sambrook, Fritsch, & Maniatis, eds., 1989).

In the alternative, chemical synthesis can be carried out in order to obtain nucleic acid probes having nucleotide sequences which correspond to N-terminal and C-terminal portions of the amino acid sequence of the polypeptide of interest. The synthesized nucleic acid probes may be used as primers in a polymerase chain reaction (PCR) carried out in accordance with recognized PCR techniques, essentially according to PCR Protocols, "A Guide to Methods and Applications", Academic Press, Michael, *et al.*, eds.,

1990, utilizing the appropriate chromosomal or cDNA library to obtain the fragment of the present invention.

One skilled in the art can readily design such probes based on the sequence disclosed herein using methods of computer alignment and sequence analysis known in the art ("Molecular Cloning: A Laboratory Manual", 1989, *supra*). The hybridization probes of the present invention can be labeled by standard labeling techniques such as with a radiolabel, enzyme label, fluorescent label, biotin-avidin label, chemiluminescence, and the like. After hybridization, the probes may be visualized using known methods.

The nucleic acid probes of the present invention include RNA, as well as DNA probes, such probes being generated using techniques known in the art. The nucleic acid probe may be immobilized on a solid support. Examples of such solid supports include, but are not limited to, plastics such as polycarbonate, complex carbohydrates such as agarose and sepharose, and acrylic resins, such as polyacrylamide and latex beads. Techniques for coupling nucleic acid probes to such solid supports are well known in the art.

The test samples suitable for nucleic acid probing methods of the present invention include, for example, cells or nucleic acid extracts of cells, or biological fluids. The samples used in the above-described methods will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts to be assayed. Methods for preparing nucleic acid extracts of cells are well known in the art and can be readily adapted in order to obtain a sample which is compatible with the method utilized.

One method of detecting the presence of nucleic acids of the invention in a sample comprises (a) contacting said sample with the above-described nucleic acid probe under conditions such that hybridization occurs, and (b) detecting the presence of said probe bound to said nucleic acid molecule. One skilled in the art would select the nucleic acid probe according to techniques known in the art as described above. Samples to be tested include but should not be limited to RNA samples of human tissue.

A kit for detecting the presence of nucleic acids of the invention in a sample comprises at least one container means having disposed therein the above-described nucleic acid probe. The kit may further comprise other containers comprising one or more of the following: wash reagents and reagents capable of detecting the presence of bound nucleic acid probe. Examples of detection reagents include, but are not limited to radiolabelled probes, enzymatic labeled probes (horseradish peroxidase, alkaline

phosphatase), and affinity labeled probes (biotin, avidin, or steptavidin). Preferably, the kit further comprises instructions for use.

In detail, a compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allow the efficient transfer of reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the probe or primers used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, and the like), and containers which contain the reagents used to detect the hybridized probe, bound antibody, amplified product, or the like. One skilled in the art will readily recognize that the nucleic acid probes described in the present invention can readily be incorporated into one of the established kit formats which are well known in the art.

DNA Constructs Comprising a BUB1-Related or NEK Kinase Nucleic Acid Molecule and Cells Containing These Constructs.

The present invention also relates to a recombinant DNA molecule comprising, 5' to 3', a promoter effective to initiate transcription in a host cell and the above-described nucleic acid molecules. In addition, the present invention relates to a recombinant DNA molecule comprising a vector and an above-described nucleic acid molecule. The present invention also relates to a nucleic acid molecule comprising a transcriptional region functional in a cell, a sequence complementary to an RNA sequence encoding an amino acid sequence corresponding to the above-described polypeptide, and a transcriptional termination region functional in said cell. The above-described molecules may be isolated and/or purified DNA molecules.

The present invention also relates to a cell or organism that contains an above-described nucleic acid molecule and thereby is capable of expressing a polypeptide. The polypeptide may be purified from cells which have been altered to express the polypeptide. A cell is said to be "altered to express a desired polypeptide" when the cell, through genetic manipulation, is made to produce a protein which it normally does not produce or which the cell normally produces at lower levels. One skilled in the art can

readily adapt procedures for introducing and expressing either genomic, cDNA, or synthetic sequences into either eukaryotic or prokaryotic cells.

A nucleic acid molecule, such as DNA, is said to be "capable of expressing" a polypeptide if it contains nucleotide sequences which contain transcriptional and translational regulatory information and such sequences are "operably linked" to nucleotide sequences which encode the polypeptide. An operable linkage is a linkage in which the regulatory DNA sequences and the DNA sequence sought to be expressed are connected in such a way as to permit gene sequence expression. The precise nature of the regulatory regions needed for gene sequence expression may vary from organism to organism, but shall in general include a promoter region which, in prokaryotes, contains both the promoter (which directs the initiation of RNA transcription) as well as the DNA sequences which, when transcribed into RNA, will signal synthesis initiation. Such regions will normally include those 5'-non-coding sequences involved with initiation of transcription and translation, such as the TATA box, capping sequence, CAAT sequence, and the like.

If desired, the non-coding region 3' to the sequence encoding a kinase of the invention may be obtained by the above-described methods. This region may be retained for its transcriptional termination regulatory sequences, such as termination and polyadenylation. Thus, by retaining the 3'-region naturally contiguous to the DNA sequence encoding a kinase of the invention, the transcriptional termination signals may be provided. Where the transcriptional termination signals are not satisfactorily functional in the expression host cell, then a 3' region functional in the host cell may be substituted.

Two DNA sequences (such as a promoter region sequence and a sequence encoding a kinase of the invention) are said to be operably linked if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region sequence to direct the transcription of a gene sequence encoding a kinase of the invention, or (3) interfere with the ability of the gene sequence of a kinase of the invention to be transcribed by the promoter region sequence. Thus, a promoter region would be operably linked to a DNA sequence if the promoter were capable of effecting transcription of that DNA sequence. Thus, to express a gene encoding a kinase of the invention, transcriptional and translational signals recognized by an appropriate host are necessary.

The present invention encompasses the expression of a gene encoding a kinase of the invention (or a functional derivative thereof) in either prokaryotic or eukaryotic cells. Prokaryotic hosts are, generally, very efficient and convenient for the production of recombinant proteins and are, therefore, one type of preferred expression system for kinases of the invention. Prokaryotes most frequently are represented by various strains of *E. coli*. However, other microbial strains may also be used, including other bacterial strains.

In prokaryotic systems, plasmid vectors that contain replication sites and control sequences derived from a species compatible with the host may be used. Examples of suitable plasmid vectors may include pBR322, pUC118, pUC119 and the like; suitable phage or bacteriophage vectors may include λgt10, λgt11 and the like; and suitable virus vectors may include pMAM-neo, pKRC and the like. Preferably, the selected vector of the present invention has the capacity to replicate in the selected host cell.

Recognized prokaryotic hosts include bacteria such as *E. coli*, *Bacillus*, *Streptomyces*, *Pseudomonas*, *Salmonella*, *Serratia*, and the like. However, under such conditions, the polypeptide will not be glycosylated. The prokaryotic host must be compatible with the replicon and control sequences in the expression plasmid.

To express a kinase of the invention (or a functional derivative thereof) in a prokaryotic cell, it is necessary to operably link the sequence encoding the kinase of the invention to a functional prokaryotic promoter. Such promoters may be either constitutive or, more preferably, regulatable (*i.e.*, inducible or derepressible). Examples of constitutive promoters include the *int* promoter of bacteriophage λ, the *bla* promoter of the β-lactamase gene sequence of pBR322, and the *cat* promoter of the chloramphenicol acetyl transferase gene sequence of pPR325, and the like. Examples of inducible prokaryotic promoters include the major right and left promoters of bacteriophage λ (P_L and P_R), the *trp*, *recA*, *λacZ*, *λacI*, and *gal* promoters of *E. coli*, the α-amylase (Ulmanen *et al.*, J. Bacteriol. 162:176-182, 1985) and the ς-28-specific promoters of *B. subtilis* (Gilman *et al.*, Gene Sequence 32:11-20, 1984), the promoters of the bacteriophages of *Bacillus* (Gryczan, In: The Molecular Biology of the Bacilli, Academic Press, Inc., NY, 1982), and *Streptomyces* promoters (Ward *et al.*, Mol. Gen. Genet. 203:468-478, 1986). Prokaryotic promoters are reviewed by Glick (Ind. Microbiot. 1:277-282, 1987), Cenatiempo (Biochimie 68:505-516, 1986), and Gottesman (Ann. Rev. Genet. 18:415-442, 1984).

Proper expression in a prokaryotic cell also requires the presence of a ribosome-binding site upstream of the gene sequence-encoding sequence. Such ribosome-binding sites are disclosed, for example, by Gold *et al.* (Ann. Rev. Microbiol. 35:365-404, 1981). The selection of control sequences, expression vectors, transformation methods, and the like, are dependent on the type of host cell used to express the gene. As used herein, "cell", "cell line", and "cell culture" may be used interchangeably and all such designations include progeny. Thus, the words "transformants" or "transformed cells" include the primary subject cell and cultures derived therefrom, without regard to the number of transfers. It is also understood that all progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. However, as defined, mutant progeny have the same functionality as that of the originally transformed cell.

Host cells which may be used in the expression systems of the present invention are not strictly limited, provided that they are suitable for use in the expression of the kinase polypeptide of interest. Suitable hosts may often include eukaryotic cells. Preferred eukaryotic hosts include, for example, yeast, fungi, insect cells, mammalian cells either *in vivo*, or in tissue culture. Mammalian cells which may be useful as hosts include HeLa cells, cells of fibroblast origin such as VERO or CHO-K1, or cells of lymphoid origin and their derivatives. Preferred mammalian host cells include SP2/0 and J558L, as well as neuroblastoma cell lines such as IMR 332, which may provide better capacities for correct post-translational processing.

In addition, plant cells are also available as hosts, and control sequences compatible with plant cells are available, such as the cauliflower mosaic virus 35S and 19S, and nopaline synthase promoter and polyadenylation signal sequences. Another preferred host is an insect cell, for example the *Drosophila* larvae. Using insect cells as hosts, the *Drosophila* alcohol dehydrogenase promoter can be used (Rubin, Science 240:1453-1459, 1988). Alternatively, baculovirus vectors can be engineered to express large amounts of kinases of the invention in insect cells (Jasny, Science 238:1653, 1987; Miller *et al.*, In: Genetic Engineering, Vol. 8, Plenum, Setlow *et al.*, eds., pp. 277-297, 1986).

Any of a series of yeast expression systems can be utilized which incorporate promoter and termination elements from the actively expressed sequences coding for glycolytic enzymes that are produced in large quantities when yeast are grown in mediums rich in glucose. Known glycolytic gene sequences can also provide very efficient

transcriptional control signals. Yeast provides substantial advantages in that it can also carry out post-translational modifications. A number of recombinant DNA strategies exist utilizing strong promoter sequences and high copy number plasmids which can be utilized for production of the desired proteins in yeast. Yeast recognizes leader sequences on cloned mammalian genes and secretes peptides bearing leader sequences (*i.e.*, pre-peptides). Several possible vector systems are available for the expression of kinases of the invention in a mammalian host.

A wide variety of transcriptional and translational regulatory sequences may be employed, depending upon the nature of the host. The transcriptional and translational regulatory signals may be derived from viral sources, such as adenovirus, bovine papilloma virus, cytomegalovirus, simian virus, or the like, where the regulatory signals are associated with a particular gene sequence which has a high level of expression. Alternatively, promoters from mammalian expression products, such as actin, collagen, myosin, and the like, may be employed. Transcriptional initiation regulatory signals may be selected which allow for repression or activation, so that expression of the gene sequences can be modulated. Of interest are regulatory signals which are temperature-sensitive so that by varying the temperature, expression can be repressed or initiated, or are subject to chemical (such as metabolite) regulation.

Expression of kinases of the invention in eukaryotic hosts requires the use of eukaryotic regulatory regions. Such regions will, in general, include a promoter region sufficient to direct the initiation of RNA synthesis. Preferred eukaryotic promoters include, for example, the promoter of the mouse metallothionein I gene sequence (Hamer *et al.*, J. Mol. Appl. Gen. 1:273-288, 1982); the TK promoter of Herpes virus (McKnight, Cell 31:355-365, 1982); the SV40 early promoter (Benoist *et al.*, Nature (London) 290:304-31, 1981); and the yeast gal4 gene sequence promoter (Johnston *et al.*, Proc. Natl. Acad. Sci. (USA) 79:6971-6975, 1982; Silver *et al.*, Proc. Natl. Acad. Sci. (USA) 81:5951-5955, 1984).

Translation of eukaryotic mRNA is initiated at the codon which encodes the first methionine. For this reason, it is preferable to ensure that the linkage between a eukaryotic promoter and a DNA sequence which encodes a kinase of the invention (or a functional derivative thereof) does not contain any intervening codons which are capable of encoding a methionine (*i.e.*, AUG). The presence of such codons results either in the formation of a fusion protein (if the AUG codon is in the same reading frame as the kinase

of the invention coding sequence) or a frame-shift mutation (if the AUG codon is not in the same reading frame as the kinase of the invention coding sequence).

A nucleic acid molecule encoding a kinase of the invention and an operably linked promoter may be introduced into a recipient prokaryotic or eukaryotic cell either as a nonreplicating DNA or RNA molecule, which may either be a linear molecule or, more preferably, a closed covalent circular molecule. Since such molecules are incapable of autonomous replication, the expression of the gene may occur through the transient expression of the introduced sequence. Alternatively, permanent expression may occur through the integration of the introduced DNA sequence into the host chromosome.

A vector may be employed which is capable of integrating the desired gene sequences into the host cell chromosome. Cells which have stably integrated the introduced DNA into their chromosomes can be selected by also introducing one or more markers which allow for selection of host cells which contain the expression vector. The marker may provide for prototrophy to an auxotrophic host, biocide resistance, *e.g.*, antibiotics, or heavy metals, such as copper, or the like. The selectable marker gene sequence can either be directly linked to the DNA gene sequences to be expressed, or introduced into the same cell by co-transfection. Additional elements may also be needed for optimal synthesis of mRNA. These elements may include splice signals, as well as transcription promoters, enhancers, and termination signals. cDNA expression vectors incorporating such elements include those described by Okayama (Mol. Cell. Biol. 3:280-289, 1983).

The introduced nucleic acid molecule can be incorporated into a plasmid or viral vector capable of autonomous replication in the recipient host. Any of a wide variety of vectors may be employed for this purpose. Factors of importance in selecting a particular plasmid or viral vector include: the ease with which recipient cells that contain the vector may be recognized and selected from those recipient cells which do not contain the vector; the number of copies of the vector which are desired in a particular host; and whether it is desirable to be able to "shuttle" the vector between host cells of different species.

Preferred prokaryotic vectors include plasmids such as those capable of replication in *E. coli* (such as, for example, pBR322, ColEl, pSC101, pACYC 184, πVX; "Molecular Cloning: A Laboratory Manual", 1989, *supra*). Bacillus plasmids include pC194, pC221, pT127, and the like (Gryczan, In: The Molecular Biology of the Bacilli, Academic Press, NY, pp. 307-329, 1982). Suitable *Streptomyces* plasmids include p1J101 (Kendall *et al.*,

J. Bacteriol. 169:4177-4183, 1987), and streptomyces bacteriophages such as ϕC31 (Chater *et al.*, In: Sixth International Symposium on Actinomycetales Biology, Akademiai Kaido, Budapest, Hungary, pp. 45-54, 1986). *Pseudomonas* plasmids are reviewed by John *et al.* (Rev. Infect. Dis. 8:693-704, 1986), and Izaki (Jpn. J. Bacteriol. 33:729-742, 1978).

Preferred eukaryotic plasmids include, for example, BPV, vaccinia, SV40, 2-micron circle, and the like, or their derivatives. Such plasmids are well known in the art (Botstein *et al.*, Miami Wntr. Symp. 19:265-274, 1982; Broach, In: "The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance", Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, p. 445-470, 1981; Broach, Cell 28:203-204, 1982; Bollon *et al.*, J. Clin. Hematol. Oncol. 10:39-48, 1980; Maniatis, In: Cell Biology: A Comprehensive Treatise, Vol. 3, Gene Sequence Expression, Academic Press, NY, pp. 563-608, 1980).

Once the vector or nucleic acid molecule containing the construct(s) has been prepared for expression, the DNA construct(s) may be introduced into an appropriate host cell by any of a variety of suitable means, *i.e.*, transformation, transfection, conjugation, protoplast fusion, electroporation, particle gun technology, calcium phosphate-precipitation, direct microinjection, and the like. After the introduction of the vector, recipient cells are grown in a selective medium, which selects for the growth of vector-containing cells. Expression of the cloned gene(s) results in the production of a kinase of the invention, or fragments thereof. This can take place in the transformed cells as such, or following the induction of these cells to differentiate (for example, by administration of bromodeoxyuracil to neuroblastoma cells or the like). A variety of incubation conditions can be used to form the peptide of the present invention. The most preferred conditions are those which mimic physiological conditions.

Antibodies, Hybridomas, Methods of Use and Kits for Detection of hBUB1-Related and NEK Kinases

The present invention relates to an antibody having binding affinity to a kinase of the invention. The polypeptide may have the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or a functional derivative thereof, or at least 9 contiguous amino acids thereof (preferably, at least 20, 30, 35, or 40 contiguous amino acids thereof).

The present invention also relates to an antibody having specific binding affinity to a kinase of the invention. Such an antibody may be isolated by comparing its binding affinity to a kinase of the invention with its binding affinity to other polypeptides. Those which bind selectively to a kinase of the invention would be chosen for use in methods requiring a distinction between a kinase of the invention and other polypeptides. Such methods could include, but should not be limited to, the analysis of altered kinase expression in tissue containing other polypeptides.

The BUB1 or NEK kinases of the present invention can be used in a variety of procedures and methods, such as for the generation of antibodies, for use in identifying pharmaceutical compositions, and for studying DNA/protein interaction.

The kinases of the present invention can be used to produce antibodies or hybridomas. One skilled in the art will recognize that if an antibody is desired, such a peptide could be generated as described herein and used as an immunogen. The antibodies of the present invention include monoclonal and polyclonal antibodies, as well fragments of these antibodies, and humanized forms. Humanized forms of the antibodies of the present invention may be generated using one of the procedures known in the art such as chimerization or CDR grafting.

The present invention also relates to a hybridoma which produces the above-described monoclonal antibody, or binding fragment thereof. A hybridoma is an immortalized cell line which is capable of secreting a specific monoclonal antibody.

In general, techniques for preparing monoclonal antibodies and hybridomas are well known in the art (Campbell, "Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology," Elsevier Science Publishers, Amsterdam, The Netherlands, 1984; St. Groth *et al.*, J. Immunol. Methods 35:1-21, 1980). Any animal (mouse, rabbit, and the like) which is known to produce antibodies can be immunized with the selected polypeptide. Methods for immunization are well known in the art. Such methods include subcutaneous or intraperitoneal injection of the polypeptide. One skilled in the art will recognize that the amount of polypeptide used for immunization will vary based on the animal which is immunized, the antigenicity of the polypeptide and the site of injection.

The polypeptide may be modified or administered in an adjuvant in order to increase the peptide antigenicity. Methods of increasing the antigenicity of a polypeptide are well known in the art. Such procedures include coupling the antigen with a

heterologous protein (such as globulin or β-galactosidase) or through the inclusion of an adjuvant during immunization.

For monoclonal antibodies, spleen cells from the immunized animals are removed, fused with myeloma cells, such as SP2/0-Agl4 myeloma cells, and allowed to become monoclonal antibody producing hybridoma cells. Any one of a number of methods well known in the art can be used to identify the hybridoma cell which produces an antibody with the desired characteristics. These include screening the hybridomas with an ELISA assay, western blot analysis, or radioimmunoassay (Lutz *et al.*, Exp. Cell Res. 175:109-124, 1988). Hybridomas secreting the desired antibodies are cloned and the class and subclass are determined using procedures known in the art (Campbell, "Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology", *supra*, 1984).

For polyclonal antibodies, antibody-containing antisera is isolated from the immunized animal and is screened for the presence of antibodies with the desired specificity using one of the above-described procedures. The above-described antibodies may be detectably labeled. Antibodies can be detectably labeled through the use of radioisotopes, affinity labels (such as biotin, avidin, and the like), enzymatic labels (such as horseradish peroxidase, alkaline phosphatase, and the like) fluorescent labels (such as FITC or rhodamine, and the like), paramagnetic atoms, and the like. Procedures for accomplishing such labeling are well-known in the art, for example, *see* Stemberger *et al.*, J. Histochem. Cytochem. 18:315, 1970; Bayer *et al.*, Meth. Enzym. 62:308, 1979; Engval *et al.*, Immunol. 109:129, 1972; Goding, J. Immunol. Meth. 13:215, 1976. The labeled antibodies of the present invention can be used for *in vitro, in vivo*, and *in situ* assays to identify cells or tissues which express a specific peptide.

The above-described antibodies may also be immobilized on a solid support. Examples of such solid supports include plastics such as polycarbonate, complex carbohydrates such as agarose and sepharose, acrylic resins such as polyacrylamide and latex beads. Techniques for coupling antibodies to such solid supports are well known in the art (Weir *et al.*, "Handbook of Experimental Immunology" 4th Ed., Blackwell Scientific Publications, Oxford, England, Chapter 10, 1986; Jacoby *et al.*, Meth. Enzym. 34, Academic Press, N.Y., 1974). The immobilized antibodies of the present invention can be used for *in vitro, in vivo*, and *in situ* assays as well as in immunochromotography.

Furthermore, one skilled in the art can readily adapt currently available procedures, as well as the techniques, methods and kits disclosed herein with regard to antibodies, to generate peptides capable of binding to a specific peptide sequence in order to generate rationally designed antipeptide peptides (Hurby *et al.*, "Application of Synthetic Peptides: Antisense Peptides", In Synthetic Peptides, A User's Guide, W.H. Freeman, NY, pp. 289-307, 1992; Kaspczak *et al.*, Biochemistry 28:9230-9238, 1989).

Anti-peptide peptides can be generated by replacing the basic amino acid residues found in the peptide sequences of the kinases of the invention with acidic residues, while maintaining hydrophobic and uncharged polar groups. For example, lysine, arginine, and/or histidine residues are replaced with aspartic acid or glutamic acid and glutamic acid residues are replaced by lysine, arginine or histidine.

The present invention also encompasses a method of detecting a BUB1-related or NEK kinase polypeptide in a sample, comprising: (a) contacting the sample with an above-described antibody, under conditions such that immunocomplexes form, and (b) detecting the presence of said antibody bound to the polypeptide. In detail, the methods comprise incubating a test sample with one or more of the antibodies of the present invention and assaying whether the antibody binds to the test sample. Altered levels of a kinase of the invention in a sample as compared to normal levels may indicate disease.

Conditions for incubating an antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the antibody used in the assay. One skilled in the art will recognize that any one of the commonly available immunological assay formats (such as radioimmunoassays, enzyme-linked immunosorbent assays, diffusion-based Ouchterlony, or rocket immunofluorescent assays) can readily be adapted to employ the antibodies of the present invention. Examples of such assays can be found in Chard ("An Introduction to Radioimmunoassay and Related Techniques" Elsevier Science Publishers, Amsterdam, The Netherlands, 1986), Bullock *et al.* ("Techniques in Immunocytochemistry," Academic Press, Orlando, FL Vol. 1, 1982; Vol. 2, 1983; Vol. 3, 1985), Tijssen ("Practice and Theory of Enzyme Immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology," Elsevier Science Publishers, Amsterdam, The Netherlands, 1985).

The immunological assay test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as blood, serum, plasma, or urine. The test samples used in the above-described method will vary based on the assay

format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can readily be adapted in order to obtain a sample which is testable with the system utilized.

A kit contains all the necessary reagents to carry out the previously described methods of detection. The kit may comprise: (i) a first container means containing an above-described antibody, and (ii) second container means containing a conjugate comprising a binding partner of the antibody and a label. In another preferred embodiment, the kit further comprises one or more other containers comprising one or more of the following: wash reagents and reagents capable of detecting the presence of bound antibodies.

Examples of detection reagents include, but are not limited to, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the chromophoric, enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. The compartmentalized kit may be as described above for nucleic acid probe kits. One skilled in the art will readily recognize that the antibodies described in the present invention can readily be incorporated into one of the established kit formats which are well known in the art.

Isolation of Compounds Which Interact With BUB1-Related or NEK Kinases

The present invention also relates to a method of detecting a compound capable of binding to a BUB1-related or NEK kinase of the invention comprising incubating the compound with a kinase of the invention and detecting the presence of the compound bound to the kinase. The compound may be present within a complex mixture, for example, serum, body fluid, or cell extracts.

The present invention also relates to a method of detecting an agonist or antagonist of kinase activity or kinase binding partner activity comprising incubating cells that produce a kinase of the invention in the presence of a compound and detecting changes in the level of kinase activity or kinase binding partner activity. The compounds thus identified would produce a change in activity indicative of the presence of the compound. The compound may be present within a complex mixture, for example, serum, body fluid, or cell extracts. Once the compound is identified it can be isolated using techniques well known in the art.

The present invention also encompasses a method of agonizing (stimulating) or antagonizing kinase associated activity in a mammal comprising administering to said mammal an agonist or antagonist to a kinase of the invention in an amount sufficient to effect said agonism or antagonism. A method of treating diseases in a mammal with an agonist or antagonist of BUB1-related or NEK kinase activity comprising administering the agonist or antagonist to a mammal in an amount sufficient to agonize or antagonize BUB1-related or NEK kinase associated functions is also encompassed in the present application.

In an effort to discover novel treatments for diseases, biomedical researchers and chemists have designed, synthesized, and tested molecules that inhibit the function of protein kinases. Some small organic molecules form a class of compounds that modulate the function of protein kinases. Examples of molecules that have been reported to inhibit the function of protein kinases include, but are not limited to, bis monocyclic, bicyclic or heterocyclic aryl compounds (PCT WO 92/20642, published November 26, 1992 by Maguire *et al.*), vinylene-azaindole derivatives (PCT WO 94/14808, published July 7, 1994 by Ballinari *et al.*), 1-cyclopropyl-4-pyridyl-quinolones (U.S. Patent No. 5,330,992), styryl compounds (U.S. Patent No. 5,217,999), styryl-substituted pyridyl compounds (U.S. Patent No. 5,302,606), certain quinazoline derivatives (EP Application No. 0 566 266 A1), seleoindoles and selenides (PCT WO 94/03427, published February 17, 1994 by Denny *et al.*), tricyclic polyhydroxylic compounds (PCT WO 92/21660, published December 10, 1992 by Dow), and benzylphosphonic acid compounds (PCT WO 91/15495, published October 17, 1991 by Dow *et al*).

Compounds that can traverse cell membranes and are resistant to acid hydrolysis are potentially advantageous as therapeutics as they can become highly bioavailable after being administered orally to patients. However, many of these protein kinase inhibitors only weakly inhibit the function of protein kinases. In addition, many inhibit a variety of protein kinases and will therefore cause multiple side-effects as therapeutics for diseases.

Some indolinone compounds, however, form classes of acid resistant and membrane permeable organic molecules. WO 96/22976 (published August 1, 1996 by Ballinari *et al.*) describes hydrosoluble indolinone compounds that harbor tetralin, naphthalene, quinoline, and indole substituents fused to the oxindole ring. These bicyclic substituents are in turn substituted with polar moieties including hydroxylated alkyl, phosphate, and ether moieties. U.S. Patent Application Serial Nos. 08/702,232, filed

August 23, 1996, entitled "Indolinone Combinatorial Libraries and Related Products and Methods for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 221/187) and 08/485,323, filed June 7, 1995, entitled "Benzylidene-Z-Indoline Compounds for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 223/298) and International Patent Publications WO 96/40116, published December 19, 1996 by Tang, *et al.*, and WO 96/22976, published August 1, 1996 by Ballinari *et al.*, all of which are incorporated herein by reference in their entirety, including any drawings, figures, or tables, describe indolinone chemical libraries of indolinone compounds harboring other bicyclic moieties as well as monocyclic moieties fused to the oxindole ring. Applications 08/702,232, filed August 23, 1996, entitled "Indolinone Combinatorial Libraries and Related Products and Methods for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 221/187), 08/485,323, filed June 7, 1995, entitled "Benzylidene-Z-Indoline Compounds for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 223/298), and WO 96/22976, published August 1, 1996 by Ballinari *et al.* teach methods of indolinone synthesis, methods of testing the biological activity of indolinone compounds in cells, and inhibition patterns of indolinone derivatives.

Other examples of substances capable of modulating kinase activity include, but are not limited to, tyrphostins, quinazolines, quinoxolines, and quinolines. The quinazolines, tyrphostins, quinolines, and quinoxolines referred to above include well known compounds such as those described in the literature. For example, representative publications describing quinazolines include Barker *et al.*, EPO Publication No. 0 520 722 A1; Jones *et al.*, U.S. Patent No. 4,447,608; Kabbe *et al.*, U.S. Patent No. 4,757,072; Kaul and Vougioukas, U.S. Patent No. 5,316,553; Kreighbaum and Comer, U.S. Patent No. 4,343,940; Pegg and Wardleworth, EPO Publication No. 0 562 734 A1; Barker *et al.*, (1991) Proc. of Am. Assoc. for Cancer Research 32:327; Bertino, J.R., (1979) Cancer Research 3:293-304; Bertino, J.R., (1979) Cancer Research 9(2 part 1):293-304; Curtin *et al.*, (1986) Br. J. Cancer 53:361-368; Fernandes *et al.*, (1983) Cancer Research 43:1117-1123 ; Ferris *et al.* J. Org. Chem. 44(2):173-178; Fry *et al.*, (1994) Science 265:1093-1095; Jackman *et al.*, (1981) Cancer Research 51:5579-5586; Jones *et al.* J. Med. Chem. 29(6):1114-1118; Lee and Skibo, (1987) Biochemistry 26(23):7355-7362; Lemus *et al.*, (1989) J. Org. Chem. 54:3511-3518; Ley and Seng, (1975) Synthesis 1975:415-522; Maxwell *et al.*, (1991) Magnetic Resonance in Medicine 17:189-196 ; Mini *et al.*, (1985) Cancer Research 45:325-330; Phillips and Castle, J. (1980) Heterocyclic

Chem. 17(19):1489-1596; Reece *et al.*, (1977) Cancer Research 47(11):2996-2999; Sculier *et al.*, (1986) Cancer Immunol. and Immunother. 23, A65; Sikora *et al.*, (1984) Cancer Letters 23:289-295; Sikora *et al.*, (1988) Analytical Biochem. 172:344-355; all of which are incorporated herein by reference in their entirety, including any drawings.

Quinoxaline is described in Kaul and Vougioukas, U.S. Patent No. 5,316,553, incorporated herein by reference in its entirety, including any drawings.

Quinolines are described in Dolle *et al.*, (1994) J. Med. Chem. 37:2627-2629; MaGuire, J. (1994) Med. Chem. 37:2129-2131; Burke *et al.*, (1993) J. Med. Chem. 36:425-432 ; and Burke *et al.* (1992) BioOrganic Med. Chem. Letters 2:1771-1774, all of which are incorporated by reference in their entirety, including any drawings.

Tyrphostins are described in Allen *et al.*, (1993) Clin. Exp. Immunol. 91:141-156; Anafi *et al.*, (1993) Blood 82:12, 3524-3529; Baker *et al.*, (1992) J. Cell Sci. 102:543-555; Bilder *et al.*, (1991) Amer. Physiol. Soc. pp. 6363-6143:C721-C730; Brunton *et al.*, (1992) Proceedings of Amer. Assoc. Cancer Rsch. 33:558; Bryckaert *et al.*, (1992) Exp. Cell Research 199:255-261; Dong *et al.*, (1993) J. Leukocyte Biology 53:53-60; Dong *et al.*, (1993) J. Immunol. 151(5):2717-2724; Gazit *et al.*, (1989) J. Med. Chem. 32, 2344-2352; Gazit *et al.*, (1993) J. Med. Chem. 36:3556-3564; Kaur *et al.*, (1994) Anti-Cancer Drugs 5:213-222; King *et al.*, (1991) Biochem. J. 275:413-418; Kuo *et al.*, (1993) Cancer Letters 74:197-202; Levitzki, A., (1992) The FASEB J. 6:3275-3282; Lyall *et al.*, (1989) J. Biol. Chem. 264:14503-14509; Peterson *et al.*, (1993) The Prostate 22:335-345; Pillemer *et al.*, (1992) Int. J. Cancer 50:80-85; Posner *et al.*, (1993) Molecular Pharmacology 45:673-683; Rendu *et al.*, (1992) Biol. Pharmacology 44(5):881-888; Sauro and Thomas, (1993) Life Sciences 53:371-376; Sauro and Thomas, (1993) J. Pharm. and Experimental Therapeutics 267(3):119-1125; Wolbring *et al.*, (1994) J. Biol. Chem. 269(36):22470-22472; and Yoneda *et al.*, (1991) Cancer Research 51:4430-4435; all of which are incorporated herein by reference in their entirety, including any drawings.

Other compounds that could be used as modulators include oxindolinones such as those described in U.S. patent application Serial No. 08/702,232 filed August 23, 1996, incorporated herein by reference in its entirety, including any drawings.

Transgenic Animals.

A variety of methods are available for the production of transgenic animals associated with this invention. DNA can be injected into the pronucleus of a fertilized egg

before fusion of the male and female pronuclei, or injected into the nucleus of an embryonic cell (*e.g.*, the nucleus of a two-cell embryo) following the initiation of cell division (Brinster *et al.*, Proc. Nat. Acad. Sci. USA 82:4438-4442, 1985). Embryos can be infected with viruses, especially retroviruses, modified to carry inorganic-ion receptor nucleotide sequences of the invention.

Pluripotent stem cells derived from the inner cell mass of the embryo and stabilized in culture can be manipulated in culture to incorporate nucleotide sequences of the invention. A transgenic animal can be produced from such cells through implantation into a blastocyst that is implanted into a foster mother and allowed to come to term. Animals suitable for transgenic experiments can be obtained from standard commercial sources such as Charles River (Wilmington, MA), Taconic (Germantown, NY), Harlan Sprague Dawley (Indianapolis, IN), etc.

The procedures for manipulation of the rodent embryo and for microinjection of DNA into the pronucleus of the zygote are well known to those of ordinary skill in the art (Hogan *et al.*, *supra*). Microinjection procedures for fish, amphibian eggs and birds are detailed in Houdebine and Chourrout (Experientia 47:897-905, 1991). Other procedures for introduction of DNA into tissues of animals are described in U.S. Patent No. 4,945,050 (Sanford *et al.*, July 30, 1990).

By way of example only, to prepare a transgenic mouse, female mice are induced to superovulate. Females are placed with males, and the mated females are sacrificed by CO_2 asphyxiation or cervical dislocation and embryos are recovered from excised oviducts. Surrounding cumulus cells are removed. Pronuclear embryos are then washed and stored until the time of injection. Randomly cycling adult female mice are paired with vasectomized males. Recipient females are mated at the same time as donor females. Embryos then are transferred surgically. The procedure for generating transgenic rats is similar to that of mice (Hammer *et al.*, Cell 63:1099-1112, 1990).

Methods for the culturing of embryonic stem (ES) cells and the subsequent production of transgenic animals by the introduction of DNA into ES cells using methods such as electroporation, calcium phosphate/DNA precipitation and direct injection also are well known to those of ordinary skill in the art (Teratocarcinomas and Embryonic Stem Cells, A Practical Approach, E.J. Robertson, ed., IRL Press, 1987).

In cases involving random gene integration, a clone containing the sequence(s) of the invention is co-transfected with a gene encoding resistance. Alternatively, the gene

encoding neomycin resistance is physically linked to the sequence(s) of the invention. Transfection and isolation of desired clones are carried out by any one of several methods well known to those of ordinary skill in the art (E.J. Robertson, *supra*).

DNA molecules introduced into ES cells can also be integrated into the chromosome through the process of homologous recombination (Capecchi, Science 244:1288-1292, 1989). Methods for positive selection of the recombination event (*i.e.*, neo resistance) and dual positive-negative selection (*i.e.*, neo resistance and gancyclovir resistance) and the subsequent identification of the desired clones by PCR have been described by Capecchi, *supra* and Joyner *et al.* (Nature 338:153-156, 1989), the teachings of which are incorporated herein in their entirety including any drawings. The final phase of the procedure is to inject targeted ES cells into blastocysts and to transfer the blastocysts into pseudopregnant females. The resulting chimeric animals are bred and the offspring are analyzed by Southern blotting to identify individuals that carry the transgene. Procedures for the production of non-rodent mammals and other animals have been discussed by others (Houdebine and Chourrout, *supra*; Pursel *et al.*, Science 244:1281-1288, 1989; and Simms *et al.*, Bio/Technology 6:179-183, 1988).

Thus, the invention provides transgenic, nonhuman mammals containing a transgene encoding a kinase of the invention or a gene affecting the expression of the kinase. Such transgenic nonhuman mammals are particularly useful as an *in vivo* test system for studying the effects of introduction of a kinase, or regulating the expression of a kinase (*i.e.*, through the introduction of additional genes, antisense nucleic acids, or ribozymes).

A "transgenic animal" is an animal having cells that contain DNA which has been artificially inserted into a cell, which DNA becomes part of the genome of the animal which develops from that cell. Preferred transgenic animals are primates, mice, rats, cows, pigs, horses, goats, sheep, dogs and cats. The transgenic DNA may encode human STE20-related kinases. Native expression in an animal may be reduced by providing an amount of antisense RNA or DNA effective to reduce expression of the receptor.

Gene Therapy

BUB1-related or NEK kinases or their genetic sequences will also be useful in gene therapy (reviewed in Miller, Nature 357:455-460, 1992). Miller states that advances have resulted in practical approaches to human gene therapy that have demonstrated

positive initial results. The basic science of gene therapy is described in Mulligan (Science 260:926-931, 1993).

In one preferred embodiment, an expression vector containing BUB1-related or NEK kinase coding sequence is inserted into cells, the cells are grown *in vitro* and then infused in large numbers into patients. In another preferred embodiment, a DNA segment containing a promoter of choice (for example a strong promoter) is transferred into cells containing an endogenous gene encoding kinases of the invention in such a manner that the promoter segment enhances expression of the endogenous kinase gene (for example, the promoter segment is transferred to the cell such that it becomes directly linked to the endogenous kinase gene).

The gene therapy may involve the use of an adenovirus containing kinase cDNA targeted to a tumor, systemic kinase increase by implantation of engineered cells, injection with kinase-encoding virus, or injection of naked kinase DNA into appropriate tissues.

Target cell populations may be modified by introducing altered forms of one or more components of the protein complexes in order to modulate the activity of such complexes. For example, by reducing or inhibiting a complex component activity within target cells, an abnormal signal transduction event(s) leading to a condition may be decreased, inhibited, or reversed. Deletion or missense mutants of a component, that retain the ability to interact with other components of the protein complexes but cannot function in signal transduction, may be used to inhibit an abnormal, deleterious signal transduction event.

Expression vectors derived from viruses such as retroviruses, vaccinia virus, adenovirus, adeno-associated virus, herpes viruses, several RNA viruses, or bovine papilloma virus, may be used for delivery of nucleotide sequences (*e.g.*, cDNA) encoding recombinant kinase of the invention protein into the targeted cell population (*e.g.*, tumor cells). Methods which are well known to those skilled in the art can be used to construct recombinant viral vectors containing coding sequences (Maniatis *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y., 1989; Ausubel *et al.*, Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley Interscience, N.Y., 1989). Alternatively, recombinant nucleic acid molecules encoding protein sequences can be used as naked DNA or in a reconstituted system *e.g.*, liposomes or other lipid systems for delivery to target cells (*e.g.*, Felgner *et al.*, Nature 337:387-8, 1989). Several other methods for the direct transfer of plasmid DNA into cells exist for

use in human gene therapy and involve targeting the DNA to receptors on cells by complexing the plasmid DNA to proteins (Miller, *supra*).

In its simplest form, gene transfer can be performed by simply injecting minute amounts of DNA into the nucleus of a cell, through a process of microinjection (Capecchi, Cell 22:479-88, 1980). Once recombinant genes are introduced into a cell, they can be recognized by the cell's normal mechanisms for transcription and translation, and a gene product will be expressed. Other methods have also been attempted for introducing DNA into larger numbers of cells. These methods include: transfection, wherein DNA is precipitated with calcium phosphate and taken into cells by pinocytosis (Chen *et al.*, Mol. Cell Biol. 7:2745-52, 1987); electroporation, wherein cells are exposed to large voltage pulses to introduce holes into the membrane (Chu *et al.*, Nucleic Acids Res. 15:1311-26, 1987); lipofection/liposome fusion, wherein DNA is packaged into lipophilic vesicles which fuse with a target cell (Felgner *et al.*, Proc. Natl. Acad. Sci. USA. 84:7413-7417, 1987); and particle bombardment using DNA bound to small projectiles (Yang *et al.*, Proc. Natl. Acad. Sci. 87:9568-9572, 1990). Another method for introducing DNA into cells is to couple the DNA to chemically modified proteins.

It has also been shown that adenovirus proteins are capable of destabilizing endosomes and enhancing the uptake of DNA into cells. The admixture of adenovirus to solutions containing DNA complexes, or the binding of DNA to polylysine covalently attached to adenovirus using protein crosslinking agents substantially improves the uptake and expression of the recombinant gene (Curiel *et al.*, Am. J. Respir. Cell. Mol. Biol., 6:247-52, 1992).

As used herein "gene transfer" means the process of introducing a foreign nucleic acid molecule into a cell. Gene transfer is commonly performed to enable the expression of a particular product encoded by the gene. The product may include a protein, polypeptide, anti-sense DNA or RNA, or enzymatically active RNA. Gene transfer can be performed in cultured cells or by direct administration into animals. Generally gene transfer involves the process of nucleic acid contact with a target cell by non-specific or receptor mediated interactions, uptake of nucleic acid into the cell through the membrane or by endocytosis, and release of nucleic acid into the cytoplasm from the plasma membrane or endosome. Expression may require, in addition, movement of the nucleic acid into the nucleus of the cell and binding to appropriate nuclear factors for transcription.

As used herein "gene therapy" is a form of gene transfer and is included within the definition of gene transfer as used herein and specifically refers to gene transfer to express a therapeutic product from a cell *in vivo* or *in vitro*. Gene transfer can be performed *ex vivo* on cells which are then transplanted into a patient, or can be performed by direct administration of the nucleic acid or nucleic acid-protein complex into the patient.

In another preferred embodiment, a vector having nucleic acid sequences encoding a BUB1-related or NEK kinase polypeptide is provided in which the nucleic acid sequence is expressed only in specific tissue. Methods of achieving tissue-specific gene expression are set forth in International Publication No. WO 93/09236, filed November 3, 1992 and published May 13, 1993.

In all of the preceding vectors set forth above, a further aspect of the invention is that the nucleic acid sequence contained in the vector may include additions, deletions or modifications to some or all of the sequence of the nucleic acid, as defined above.

In another preferred embodiment, a method of gene replacement is set forth. "Gene replacement" as used herein means supplying a nucleic acid sequence which is capable of being expressed *in vivo* in an animal and thereby providing or augmenting the function of an endogenous gene which is missing or defective in the animal.

Pharmaceutical Formulations And Routes Of Administration

The compounds described herein can be administered to a human patient *per se*, or in pharmaceutical compositions where it is mixed with other active ingredients, as in combination therapy, or suitable carriers or excipient(s). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition.

A. Routes Of Administration

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intravenous, intramedullary injections, as well as intrathecal, direct intraventricular, intraperitoneal, intranasal, or intraocular injections.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a solid tumor, often in a depot or sustained release formulation.

Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with tumor-specific antibody. The liposomes will be targeted to and taken up selectively by the tumor.

B. Composition/Formulation

The pharmaceutical compositions of the present invention may be manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes.

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Proper formulation is dependent upon the route of administration chosen.

For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Suitable carriers include excipients such as, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl- cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate.

Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to

the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration.

For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of *e.g.* gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain

suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a cosolvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The cosolvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:D5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, *e.g.* polyvinyl pyrrolidone; and other sugars or polysaccharides may substitute for dextrose.

Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as

semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols.

Many of the tyrosine or serine/threonine kinase modulating compounds of the invention may be provided as salts with pharmaceutically compatible counterions. Pharmaceutically compatible salts may be formed with many acids, including but not limited to hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents that are the corresponding free base forms.

C. Effective Dosage

Pharmaceutical compositions suitable for use in the present invention include compositions where the active ingredients are contained in an amount effective to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount of compound effective to prevent, alleviate or ameliorate symptoms of disease or prolong the survival of the subject being treated. Determination of a therapeutically effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein.

For any compound used in the methods of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC_{50} as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the tyrosine or serine/threonine kinase activity). Such information can be used to more accurately determine useful doses in humans.

Toxicity and therapeutic efficacy of the compounds described herein can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the

ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD_{50} and ED_{50}. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. (See *e.g.*, Fingl *et al.*, 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1).

Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the kinase modulating effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from *in vitro* data; *e.g.*, the concentration necessary to achieve 50-90% inhibition of the kinase using the assays described herein. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%.

In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

D. Packaging

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. The pack or dispenser may also be accompanied with a notice associated with the container in form prescribed by a governmental agency regulating the manufacture, use, or sale of

pharmaceuticals, which notice is reflective of approval by the agency of the form of the polynucleotide for human or veterinary administration. Such notice, for example, may be the labeling approved by the U.S. Food and Drug Administration for prescription drugs, or the approved product insert. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition. Suitable conditions indicated on the label may include treatment of a tumor, inhibition of angiogenesis, treatment of fibrosis, diabetes, and the like.

Functional Derivatives

Also provided herein are functional derivatives of a polypeptide or nucleic acid of the invention. By "functional derivative" is meant a "chemical derivative," "fragment," or "variant," of the polypeptide or nucleic acid of the invention, which terms are defined below. A functional derivative retains at least a portion of the function of the protein, for example reactivity with an antibody specific for the protein, enzymatic activity or binding activity mediated through noncatalytic domains, which permits its utility in accordance with the present invention. It is well known in the art that due to the degeneracy of the genetic code numerous different nucleic acid sequences can code for the same amino acid sequence. Equally, it is also well known in the art that conservative changes in amino acid can be made to arrive at a protein or polypeptide that retains the functionality of the original. In both cases, all permutations are intended to be covered by this disclosure.

Included within the scope of this invention are the functional equivalents of the herein-described isolated nucleic acid molecules. The degeneracy of the genetic code permits substitution of certain codons by other codons that specify the same amino acid and hence would give rise to the same protein. The nucleic acid sequence can vary substantially since, with the exception of methionine and tryptophan, the known amino acids can be coded for by more than one codon. Thus, portions or all of the BUB1 and NEK genes could be synthesized to give a nucleic acid sequence significantly different from that shown in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, or SEQ ID NO:9. The encoded amino acid sequence thereof would, however, be preserved.

In addition, the nucleic acid sequence may comprise a nucleotide sequence which results from the addition, deletion or substitution of at least one nucleotide to the 5'-end and/or the 3'-end of the nucleic acid formula shown in SEQ ID NO:1, SEQ ID NO:3, SEQ

ID NO:5, SEQ ID NO:7, or SEQ ID NO:9, or a derivative thereof. Any nucleotide or polynucleotide may be used in this regard, provided that its addition, deletion or substitution does not alter the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10 which is encoded by the nucleotide sequence. For example, the present invention is intended to include any nucleic acid sequence resulting from the addition of ATG as an initiation codon at the 5'-end of the inventive nucleic acid sequence or its derivative, or from the addition of TTA, TAG or TGA as a termination codon at the 3'-end of the inventive nucleotide sequence or its derivative. Moreover, the nucleic acid molecule of the present invention may, as necessary, have restriction endonuclease recognition sites added to its 5'-end and/or 3'-end.

Such functional alterations of a given nucleic acid sequence afford an opportunity to promote secretion and/or processing of heterologous proteins encoded by foreign nucleic acid sequences fused thereto. All variations of the nucleotide sequence of the BUB1 and NEK genes of the invention and fragments thereof permitted by the genetic code are, therefore, included in this invention.

Further, it is possible to delete codons or to substitute one or more codons with codons other than degenerate codons to produce a structurally modified polypeptide, but one which has substantially the same utility or activity as the polypeptide produced by the unmodified nucleic acid molecule. As recognized in the art, the two polypeptides are functionally equivalent, as are the two nucleic acid molecules that give rise to their production, even though the differences between the nucleic acid molecules are not related to the degeneracy of the genetic code.

A "chemical derivative" of the complex contains additional chemical moieties not normally a part of the protein. Covalent modifications of the protein or peptides are included within the scope of this invention. Such modifications may be introduced into the molecule by reacting targeted amino acid residues of the peptide with an organic derivatizing agent that is capable of reacting with selected side chains or terminal residues, as described below.

Cysteinyl residues most commonly are reacted with alpha-haloacetates (and corresponding amines), such as chloroacetic acid or chloroacetamide, to give carboxymethyl or carboxyamidomethyl derivatives. Cysteinyl residues also are derivatized by reaction with bromotrifluoroacetone, chloroacetyl phosphate, N-alkylmaleimides, 3-nitro-2-pyridyl disulfide, methyl 2-pyridyl disulfide, p-

chloromercuribenzoate, 2-chloromercuri-4-nitrophenol, or chloro-7-nitrobenzo-2-oxa-1,3-diazole.

Histidyl residues are derivatized by reaction with diethylprocarbonate at pH 5.5-7.0 because this agent is relatively specific for the histidyl side chain. Para-bromophenacyl bromide also is useful; the reaction is preferably performed in 0.1 M sodium cacodylate at pH 6.0.

Lysinyl and amino terminal residues are reacted with succinic or other carboxylic acid anhydrides. Derivatization with these agents has the effect or reversing the charge of the lysinyl residues. Other suitable reagents for derivatizing primary amine containing residues include imidoesters such as methyl picolinimidate; pyridoxal phosphate; pyridoxal; chloroborohydride; trinitrobenzenesulfonic acid; O-methylisourea; 2,4 pentanedione; and transaminase-catalyzed reaction with glyoxylate.

Arginyl residues are modified by reaction with one or several conventional reagents, among them phenylglyoxal, 2,3-butanedione, 1,2-cyclohexanedione, and ninhydrin. Derivatization of arginine residues requires that the reaction be performed in alkaline conditions because of the high pK_a of the guanidine functional group. Furthermore, these reagents may react with the groups of lysine as well as the arginine alpha-amino group.

Tyrosyl residues are well-known targets of modification for introduction of spectral labels by reaction with aromatic diazonium compounds or tetranitromethane. Most commonly, N-acetylimidizol and tetranitromethane are used to form O-acetyl tyrosyl species and 3-nitro derivatives, respectively.

Carboxyl side groups (aspartyl or glutamyl) are selectively modified by reaction with carbodiimide (R'-N-C-N-R') such as 1-cyclohexyl-3-(2-morpholinyl(4-ethyl) carbodiimide or 1-ethyl-3-(4-azonia-4,4-dimethylpentyl) carbodiimide. Furthermore, aspartyl and glutamyl residues are converted to asparaginyl and glutaminyl residues by reaction with ammonium ions.

Glutaminyl and asparaginyl residues are frequently deamidated to the corresponding glutamyl and aspartyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Either form of these residues falls within the scope of this invention.

Derivatization with bifunctional agents is useful, for example, for cross-linking the component peptides of the protein to each other or to other proteins in a complex to a

water-insoluble support matrix or to other macromolecular carriers. Commonly used cross-linking agents include, for example, 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), and bifunctional maleimides such as bis-N-maleimido-1,8-octane. Derivatizing agents such as methyl-3-[p-azidophenyl) dithiolpropioimidate yield photoactivatable intermediates that are capable of forming crosslinks in the presence of light. Alternatively, reactive water-insoluble matrices such as cyanogen bromide-activated carbohydrates and the reactive substrates described in U.S. Patent Nos. 3,969,287; 3,691,016; 4,195,128; 4,247,642; 4,229,537; and 4,330,440 are employed for protein immobilization.

Other modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the alpha-amino groups of lysine, arginine, and histidine side chains (Creighton, T.E., Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)), acetylation of the N-terminal amine, and, in some instances, amidation of the C-terminal carboxyl groups.

Such derivatized moieties may improve the stability, solubility, absorption, biological half life, and the like. The moieties may alternatively eliminate or attenuate any undesirable side effect of the protein complex and the like. Moieties capable of mediating such effects are disclosed, for example, in Remington's Pharmaceutical Sciences, 18th ed., Mack Publishing Co., Easton, PA (1990).

The term "fragment" is used to indicate a polypeptide derived from the amino acid sequence of the proteins, of the complexes having a length less than the full-length polypeptide from which it has been derived. Such a fragment may, for example, be produced by proteolytic cleavage of the full-length protein. Preferably, the fragment is obtained recombinantly by appropriately modifying the DNA sequence encoding the proteins to delete one or more amino acids at one or more sites of the C-terminus, N-terminus, and/or within the native sequence. Fragments of a protein are useful for screening for substances that act to modulate signal transduction, as described herein. It is understood that such fragments may retain one or more characterizing portions of the native complex. Examples of such retained characteristics include: catalytic activity; substrate specificity; interaction with other molecules in the intact cell; regulatory

functions; or binding with an antibody specific for the native complex, or an epitope thereof.

Another functional derivative intended to be within the scope of the present invention is a "variant" polypeptide which either lacks one or more amino acids or contains additional or substituted amino acids relative to the native polypeptide. The variant may be derived from a naturally occurring complex component by appropriately modifying the protein DNA coding sequence to add, remove, and/or to modify codons for one or more amino acids at one or more sites of the C-terminus, N-terminus, and/or within the native sequence. It is understood that such variants having added, substituted and/or additional amino acids retain one or more characterizing portions of the native protein, as described above.

A functional derivative of a protein with deleted, inserted and/or substituted amino acid residues may be prepared using standard techniques well-known to those of ordinary skill in the art. For example, the modified components of the functional derivatives may be produced using site-directed mutagenesis techniques (as exemplified by Adelman *et al.*, 1983, DNA 2:183) wherein nucleotides in the DNA coding the sequence are modified such that a modified coding sequence is modified, and thereafter expressing this recombinant DNA in a prokaryotic or eukaryotic host cell, using techniques such as those described above. Alternatively, proteins with amino acid deletions, insertions and/or substitutions may be conveniently prepared by direct chemical synthesis, using methods well-known in the art. The functional derivatives of the proteins typically exhibit the same qualitative biological activity as the native proteins.

EXAMPLES

The examples below are not limiting and are merely representative of various aspects and features of the present invention. The examples below demonstrate the isolation and characterization of the serine/threonine kinases of the invention.

EXAMPLE 1: Isolation of cDNAs Encoding Human BUB1 Protein Kinase

A. Materials and Methods

Identification of novel clones

Total RNA was isolated using the Guanidine Salts/Phenol extraction protocol of Chomczynski and Sacchi (P. Chomczynski and N. Sacchi, Anal. Biochem. 162, 156 (1987)) from primary human colon cancer cell line HCC2998. This RNA was used to generate single-stranded cDNA using the Superscript Preamplification System (GIBCO BRL, Gaithersburg, MD; Gerard, GF *et al.* (1989), FOCUS 11, 66) under conditions recommended by the manufacturer. A typical reaction used 10 μg total RNA with 1.5 μg oligo$(dT)_{12-18}$ in a reaction volume of 60 μL. The product was treated with RNaseH and diluted to 100 μL with H_20. For subsequent PCR amplification, 3 μL of this single stranded cDNA was used in each reaction.

Degenerate oligonucleotides were synthesized on an Applied Biosystems 3948 DNA synthesizer using established phosphoramidite chemistry, precipitated with ethanol, and used unpurified for PCR. The sequence of the degenerate oligonucleotide primers is as follows:

BUB HRD = 5'- GAGGTGGGCATHATHCAYGGNA-3' (SEQ ID NO:14).

BUB DVW= 5'- GATGAATTRTCNGGYTTNAKRTCNCC-3'(SEQ ID NO:15).

These primers were derived from the sense and antisense strands of amino acid motifs BUB HRD = EVGIIHG (SEQ ID NO:16) and BUB DVW = GDXKPDNS (SEQ ID NO:17), respectively. Degenerate nucleotide residue designations are: N = A, C, G, or T; R = A or G; Y = C or T; H = A, C or T not G; and D = A, G or T not C.

PCR reactions were performed using degenerate primers applied to the NCI colon HCC 2998 single-stranded cDNA. The primers were added at a final concentration of 25 μM each to a mixture containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM each deoxynucleoside triphosphate, 0.001% gelatin, 1.5 U AmpliTaq DNA Polymerase (Perkin-Elmer/Cetus), and 3 μL cDNA. Following a 3 min denaturation at 95 °C, the cycling conditions were 94 °C for 30 seconds, 37 °C for 1 min, and 72 °C for 1 min 45 seconds for 3 cycles, followed by 94 °C for 30 seconds, 50 °C for 1 min, and 72 °C for 1 min 45 seconds for 35 cycles. PCR fragments migrating between 220 to 250 bp were isolated from 2% agarose gels using a GeneClean Kit (Bio101) and T-A cloned into the pCRII vector (Invitrogen Corp. U.S.A.) according to the manufacturer's protocol.

Colonies were selected for plasmid DNA mini-preparations, the plasmid DNA was purified using Qiagen columns, and was sequenced using a cycle sequencing dye-terminator kit with AmpliTaq DNA Polymerase, FS (ABI, Foster City, CA). Sequencing reaction products were run on an ABI Prism 377 DNA Sequencer and analyzed using the BLAST alignment algorithm (Altschul, S.F. *et al.* (1990) J.MolBiol. 215:403-10).

B. Isolation of cDNA Clones

A human Leukemia 5'-Stretch Plus λgt11 cDNA library (Clontech) was probed with the 238 bp PCR fragment of SuSTK34a (R51-40-18) corresponding to human BUB1. Probes were ^{32}P-labeled by random priming and used at $2x10^6$ cpm/mL following standard techniques for library screening. Pre-hybridization (3 h) and hybridization (overnight) were conducted at 42 °C in 5X SSC, 5X Denhart's solution, 2.5% dextran sulfate, 50 mM $Na_2PO_4/NaHPO_4$, pH 7.0, 50% formamide with 100 mg/mL denatured salmon sperm DNA. Stringent washes were performed at 65 °C in 0.1X SSC and 0.1% SDS. DNA sequencing was carried out on both strands using a cycle sequencing dye-terminator kit with AmpliTaq DNA Polymerase, FS (ABI, Foster City, CA). Sequencing reaction products were run on an ABI Prism 377 DNA Sequencer.

Results

Degenerate PCR cloning lead to the isolation of two fragments corresponding to two isoforms of human BUB1 named SuSTK34a (R51-40-18), and SuSTK34b (R51-40-4). SuSTK34a and SuSTK34b are identical except for the deletion of 15 nucleotides within kinase subdomain VII of SuSTK34b, resulting in a predicted 5 amino acid deletion.

The full length human *bub*1 clone was isolated as four EcoRI cDNA fragments: J6E0.45#9, J7E1.3#3, J6E1.3#5, and J2E0.9#8. In addition, a 3436 bp *Not*I fragment of a single λ clone was subcloned into a pBluescript(SKII+) vector, giving rise to pSG110. The full length human *bub*1 cDNA (SEQ ID N0.1) is 3408 bp long and consists of a 3255 bp ORF (45-3300) flanked by a 44 bp (1-44) 5' UTR and a 108 bp 3' UTR. A potential polyadenylation signal begins at position 3404. The sequence flanking the ATG conforms to the Kozak consensus for an initiating methionine, and is believed to be the translational start site for the human BUB1.

EXAMPLE 2: Expression Analysis of Mammalian BUB1 Protein Kinase

Materials and Methods

Northern Blot Analysis

The expression of *bub1* was analyzed by Northern blots prepared by running 20 µg of total RNA from 22 human adult tissues (thymus, lung, duodenum, colon, testis, brain, cerebellum, cortex, salivary gland, liver, pancreas, kidney, spleen, stomach, uterus, prostate, skeletal muscle, placenta, mammary gland, bladder, lymph node, adipose tissue), 2 human fetal normal tissues (fetal liver, fetal brain), and 60 human tumor cell lines (HOP-92, EKVX, NCI-H23, NCI-H226, NCI-H322M, NCI-H460, NCI-H522, A549, HOP-62, OVCAR-3, OVCAR-4, OVCAR-5, OVCAR-8, IGROV1, SK-OV-3, SNB-19, SNB-75, U251, SF-268, SF-295, SF-539, CCRF-CEM, K-562, MOLT-4, HL-60, RPMI 8226, SR, DU-145, PC-3, HT-29, HCC-2998, HCT-116, SW620, Colo 205, HTC15, KM-12, UO-31, SN12C, A498, CaKi1, RXF-393, ACHN, 786-0, TK-10, LOX IMVI, Malme-3M, SK-MEL-2, SK-MEL-5, SK-MEL-28, UACC-62, UACC-257, M14, MCF-7, MCF-7/ADR RES, Hs578T, MDA-MB-231, MDA-MB-435, MDA-N, BT-549, T47D).

The total RNA samples were run on denaturing formaldehyde 1% agarose gels and transferred onto nitrocellulose membranes (BioRad, CA). Two additional human normal tissue Northern blots containing 2 µg polyA+ mRNA per lane from 13 different human normal tissues (heart, brain, placenta, lung, liver, kidney, pancreas, spleen, lymph node, thymus, peripheral blood leukocyte, bone marrow, fetal liver) and 8 human tumor cell lines (HL-60, Hela Cell S3, K-562, MOLT-4, Raji, SW480, A549, G361) on a charge-modified nylon membrane (human blot #7760-1, human immune system blot #7754-1, and human cancer cell line blot #7757-1, Clontech, Palo Alto, CA) were also hybridized. Filters were hybridized with randomly primed [α^{32}P]dCTP-labeled probes synthesized from the *Eco*RI/*Pst*I 809 bp of human *bub1* fragment. Hybridization was performed at 42 °C overnight in 6X SSC, 0.1% SDS, 1X Denhardt's solution, 100 µg/mL denatured herring sperm DNA with 1-2 x 10^6 cpm/mL of ^{32}P-labeled DNA probes. The filters were washed in 0.1X SSC/0.1% SDS, 65 °C, and exposed on a Molecular Dynamics phosphorimager.

STK plasmid array blots were prepared by loading 0.5 µg of denatured plasmid encoding numerous STKs, including *bub1* (pSG110), onto a nylon membrane. The [α^{32}P]dCTP labeled single stranded DNA probes were synthesized from RNA isolated

from immune cells or tissues. Hybridization was performed at 42 °C for 16 hours in 6X SSC, 0.1% SDS, 1X Denhardt's solution, 100 μg/mL denatured herring sperm DNA with 10^6 cpm/mL of [α^{32}P]dCTP labeled single stranded probe. The filters were washed in 0.1X SSC/0.1% SDS, 65 °C, and exposed for quantitative analysis on a Molecular Dynamics phosphorimager.

Results

Expression of Human BUB1 RNA in Normal Tissues and Tumor Cell Lines

A single human *bub1* mRNA transcript of approximately 4.2 kb was identified by Northern analysis, and was found to be expressed in nine human normal tissues (thymus, testis, heart, placenta, lung, lymph node, thymus, bone marrow, fetal liver). The remainder of the 19 human normal tissues (duodenum, colon, brain, cerebellum, cortex, salivary gland, liver, pancreas, kidney, spleen, stomach, uterus, prostate, skeletal muscle, fetal brain, mammary gland, bladder, adipose tissue, peripheral blood leukocyte) were negative. Of the human tumor cell lines, 60 of 64 were positive for expression of human *bub1*. Very strong expression of *bub1* was detected in the following cell lines: K-562, SW480, NCI-H23, NCI-H226, NCI-H522, OVCAR-4, CCRF-CEM, RPMI 8226, HTC15, CaKi1, RXF-393, 786-0, TK-10, SK-MEL-28, UACC-62, M14, MCF-7, MCF-7/ADR RES, MDA-MB-231, MDA-N. BUB1 expression was negative in four cell lines (SK-OV3, PC-3, HT-29, A498), and was moderate in the remaining cell lines (Table 1).

TABLE 1

bub1 RNA Expression Analysis
in Human Normal Tissues and Cancer Cells

Cell type	Origin	BUB1
Thymus	Normal tissue	3
Fetal liver	Normal tissue	2
Lung	Normal tissue	0
Duodenum	Normal tissue	0
Colon	Normal tissue	0
Testis	Normal tissue	4
Brain	Normal tissue	0
Cerebellum	Normal tissue	0
Cortex	Normal tissue	0
Salivary gland	Normal tissue	0
Liver	Normal tissue	0

Pancreas	Normal tissue	0
Kidney	Normal tissue	0
Spleen	Normal tissue	0
Stomach	Normal tissue	0
Uterus	Normal tissue	0
Prostate	Normal tissue	0
Skeletal muscle	Normal tissue	0
Fetal brain	Normal tissue	0
Placenta	Normal tissue	0
Mammary gland	Normal tissue	0
Bladder	Normal tissue	0
Lymph node	Normal tissue	0
Adipose tissue	Normal tissue	0
Heart	Normal tissue	2*
Brain	Normal tissue	0*
Placenta	Normal tissue	1*
Lung	Normal tissue	1*
Liver	Normal tissue	0*
Kidney	Normal tissue	0*
Pancreas	Normal tissue	0*
Spleen	Normal tissue	0*
Lymph node	Normal tissue	1*
Thymus	Normal tissue	1*
PBL	Normal tissue	0*
Bone marrow	Normal tissue	1*
Fetal liver	Normal tissue	1*
HL-60	Leukemia	1*
Hela	Cervical tumor	2*
K-562	Leukemia	4*
MOLT-4	Leukemia	1*
Raji	Lymphoma	1*
SW480	Colon tumor	3*
A549	Lung tumor	1*
G361	Melanoma	1*
HOP-92	Lung tumor	2
EKVX	Lung tumor	2

NCI-H23	Lung tumor	3
NCI-H226	Lung tumor	3
NCI-H322M	Lung tumor	2
NCI-H460	Lung tumor	1
NCI-H522	Lung tumor	3
A549	Lung tumor	0
HOP-62	Lung tumor	2
OVCAR-3	Ovarian tumor	1
OVCAR-4	Ovarian tumor	4
OVCAR-5	Ovarian tumor	2
OVCAR-8	Ovarian tumor	2
IGROV1	Ovarian tumor	2
SK-OV-3	Ovarian tumor	0
SNB-19	CNS tumor	1
SNB-75	CNS tumor	1
U251	CNS tumor	1
SF-268	CNS tumor	1
SF-295	CNS tumor	2
SF-539	CNS tumor	2
CCRF-CEM	Leukemia	3
K-562	Leukemia	2
MOLT-4	Leukemia	2
HL-60	Leukemia	0
RPMI 8226	Leukemia	4
SR	Leukemia	2
DU-145	Prostate	2
PC-3	Prostate	0
HT-29	Colon tumor	0
HCC-2998	Colon tumor	2
HCT-116	Colon tumor	2
SW620	Colon tumor	2
Colo 205	Colon tumor	2
HTC15	Colon tumor	3
KM-12	Colon tumor	1
UO-31	Colon tumor	1
SN12C	Kidney tumor	1
A498	Kidney tumor	0
CaKi1	Kidney tumor	3
RXF-393	Kidney tumor	3
ACHN	Kidney tumor	1
786-0	Kidney tumor	3

TK-10	Kidney tumor	4
LOX IMVI	Melanoma	2
Malme-3M	Melanoma	1
SK-MEL-2	Melanoma	2
SK-MEL-5	Melanoma	2
SK-MEL-28	Melanoma	4
UACC-62	Melanoma	4
UACC-257	Melanoma	1
M14	Melanoma	3
MCF-7	Breast tumor	3
MCF-7/ADR RES	Breast tumor	3
Hs578T	Breast tumor	2
MDA-MB-231	Breast tumor	3
MDA-MB-435	Breast tumor	2
MDA-N	Breast tumor	3
BT-549	Breast tumor	2
T47D	Breast tumor	2

Scale: 0=Negative, 1=weak, 2=medium,
3=strong, 4=very strong
* mRNA Northern blot

Expression of Human *bub1* in Immune Tissues and Cells

Human *bub1* expression was analyzed by hybridization of [α^{32}P]dCTP-labeled single stranded DNA from human immune cells to a nylon membrane containing plasmid DNA encoding *bub1* and other STKs. This DNA array-based expression analysis was performed on 14 human immune tissues and cell lines: thymus, dendrocytes, mast cells, monocytes, B cells (primary, Jurkat, RPMI8226, SR), T cells (CD8/CD4+, TH1, TH2, CEM, MOLT4), and K562 (megakaryocytes). Expression of *bub1* was restricted to 4 of these immune sources, RPMI (B cells), MOLT4 (T cells), CEM (T cells), and primary monocytes, and was absent from the other 10 samples (data not shown). This expression profile indicates that human *bub1* is expressed in several hematopoietic cell types.

EXAMPLE 3: Human *bub1* Gene Expression Plasmid Construction and Recombinant *bub1* Expression

Results

Expression Plasmid Construction

Expression constructs were generated for human *bub1* including: a) pcDNA3.1 expression plasmids containing: the full length *bub1* nucleotide base sequence with and without an HA tag; the nucleotide base sequence encoding the N-terminal domain (amino acid 1 to 332) of BUB1 with and without an HA tag; the nucleotide base sequence encoding the central-domain (amino acid 333 to 757) of BUB1 with and without an HA tag; the nucleotide base sequence encoding the C-terminal kinase domain (amino acid 775 to 1085) of BUB1 with and without an HA tag, and the nucleotide base sequence of full length *bub1* with a Lys to Arg mutation at amino acid 821 of BUB1 with and without an HA tag; b) pGEX 4T-3 fusion plasmids containing: the full length nucleotide base sequence of *bub1*; the nucleotide base sequence encoding the N-terminal domain (amino acid 1 to 332) of BUB1; the nucleotide base sequence encoding the central domain (amino acid 333 to 757) of BUB1; the nucleotide base sequence encoding the C-terminal kinase domain (amino acid 775 to 1085) of BUB1; and c) pEGFP fusion plasmid containing the full length nucleotide base sequence encoding *bub1*. The BUB1 N-terminal domain and the K to R mutant are predicted to function as dominant negative constructs, and will be used to elucidate the function of BUB1.

EXAMPLE 4: Recombinant Expression of BUB1 Serine/threonine Protein Kinase

Materials and Methods

Transient Expression of BUB1 in Mammalian Cells

The pcDNA expression plasmids (10 μg DNA/100 mm plate) containing sequences encoding the full length, truncated, or the kinase dead form of BUB1, were introduced into human 293 cells with lipofectamine (Gibco BRL). After 72 hours, the cells were harvested in 0.5 mL solubilization buffer (20 mM HEPES pH 7.4, 150 mM NaCl, 10% glycerol, 1% Triton X-100, 1 mM EGTA, 2 mM phenylmethylsulfonyl fluoride, 1 μg/mL aprotinin). Sample aliquots were resolved by SDS polyacrylamide gel electrophoresis (PAGE) on 7.5% acrylamide/0.5% bis-acrylamide gels and electrophoretically transferred to nitrocellulose. Non-specific binding was blocked by pre-incubating blots in Blotto (phosphate buffered saline containing 5% w/v non-fat dried

milk and 0.2% v/v nonidet P-40 (Sigma)). Recombinant protein was detected using various BUB1 antisera, or anti-HA specific antisera, by Western analysis. BUB1-antisera was diluted 1:1,000-fold, and anti-HA monoclonal antibody (Babco, Berkeley, California) was diluted 1:10,000-fold for the primary antibody reactions of Western blotting. Goat anti-rabbit HRP (BioRad, Richmond, California) was used as a secondary antibody for the BUB1-antisera, and goat anti-mouse HRP (BioRad, Richmond, California) was used as a secondary antibody for the anti-HA antibody. ECL reagent (Amersham Life Science,) was used for detection.

Results

Cell lysates (30 μg each) were applied to 7.5% SDS PAGE followed by Western blotting. The BUB1 protein has a predicted molecular mass of 120 kDa. Recombinant human BUB1 migrates as a single polypeptide band of 150 kDa. The higher than predicted molecular mass is likely due to post-translational modification, possibly serine/threonine phosphorylation. The truncated forms migrated as predicted: the N-terminal domain at 45 kDa (with HA tag); the central-domain at 56 kDa.

EXAMPLE 5: Generation of Specific Immunoreagents to BUB1 Serine/threonine Protein Kinase

Materials and Methods

Specific immunoreagents were raised in rabbits against MAP-conjugated synthetic peptides corresponding to human BUB1. Each peptide was injected into two rabbits. The peptide number, amino acid sequence, and the region of BUB1 from which the peptide was derived is as follows:

505A: QKYNQRRKHEQWVNE (SEQ ID NO:18; aa257-271; N-terminal); 506A: DDKDEWQSLDQNED (SEQ ID NO:19; aa485-498; Central domain); 507A: KNIQKFVLKVQK (SEQ ID NO:20; aa813-824; catalytic domain); and 508A: RQKLKKVFQQHYTNK (SEQ ID NO:21; aa1051-1065; catalytic domain).

Additional immunoreagents were generated by immunizing rabbits with the bacterially expressed GST-fusion proteins containing the N-terminal or C-terminal domain of BUB1.

Affinity Purification Of Anti-BUB1 Peptide Antibody

Ten mg BUB1 peptide was covalently coupled to 2 mL Immunolink (Pierce AminoLink Plus). The link consists of an aldehyde support arm that forms a covalent

bond to primary amines via a Schiff base intermediate. The BUB1 peptide (ligand) was immobilized to AminoLink Plus coupling gel. The remaining active sites were blocked by adding the reducing agents. The column was then washed prior to use.

Twenty mL of filtered anti-BUB1 antiserum was applied to the immobilized antigen column. The purified anti-BUB1 IgG was eluted with 1 M propionic acid, neutralized with 3 M Tris, pH 7.4 while eluting, and buffered with 100 mM TEA buffered with 1 M Tris pH 6.8. The purified antibody was dialyzed against phosphate buffered saline pH 7.4, and was stored in 20% glycerol and 0.03% NaN_3 in PBS.

Results

Antisera generated from the anti-BUB1 peptides 506A, 507A, and 508A, were tested by Western analysis using the GST-BUB1 fusion proteins. Binding activity was clearly detected from the bleeds from both rabbits, but 505A failed to elicit a specific immunogenic response.

The affinity purified anti-BUB1 antibody was tested by Western analysis using the GST-BUB1 fusion protein, or the mammalian cell lysates from transiently transfected and expressed BUB1 protein. The purified anti-506A was very antigenic, and could be diluted 1:20,000-fold for testing against the recombinant BUB1 protein. Anti-507A and anti-508A can be purified similarly.

EXAMPLE 6: Endogenous Expression of BUB1Serine/threonine Protein Kinase

Materials and Methods

Multiple human tumor cell lines (HT-29, Calu-6, HCT15, MCF7, A549, C33a, HeLa, HT-29, H1299, SH422C, COLO205, SW480, SW837, SW948, and SNB19) were grown to confluence. Protein lysates were isolated in HNTG buffer (20 mM HEPES, pH 7.4, 150 mM NaCl, 10% glycerol, 1% Triton X-100, 1 mM EGTA, 2 mM phenylmethylsulfonylfluoride, 1 μg/mL aprotinin). The protein concentration was determined by BioRad assay (BioRad, Richmond, California). Protein from each sample (30 μg each) was tested by Western analysis with the purified anti-BUB1 506A antibody.

Results

Fifteen of the human tumor cell lines expressed detectable levels of endogenous BUB1 protein. However, extremely low RNA and protein levels of BUB1 were found in MCF7 and SNB19 cell lines.

EXAMPLE 7: Cell Cycle Regulation of BUB1

Materials and Methods

HeLa cells were synchronized by adding 0.5 M thymidine and 250 μg/mL aphidicolin. Cells were harvested at 0, 2, 4, 6, 8, 9, 10, 11, and 12 hours after thymidine and aphidicolin were removed. Other experiments were performed by blocking HeLa cells at different cell cycle stages by serum starvation, with thymidine (G1/S/G2), or with nocodozole (M). Total RNA was isolated from the synchronized cells and cell cycle blocked cells. Total RNA (10 μg) was loaded on the nylon membranes. The RNA blots were hybridized with a *bub1 Eco*RI/*Pst*I 809 bp probe.

Results

Northern blots containing RNA derived from HeLa cells either in an exponential growth state or arrested in various stages of the cell cycle were probed for expression of *bub1*. A detectable *bub1* signal was observed in the exponentially growing cells, whereas no signal was present in samples from cells grown under serum starvation or blocked in the G_1/S/G_2 stages by thymidine/aphidicolin. However, high levels of *bub1* expression were observed in cells arrested in mitosis by nocodazole. The *bub1* RNA was expressed at very low levels immediately following release from synchrony, and peaked at 8 to 10 hours (period of mitotic division) after drug release. The expression levels were decreased at 11 to 12 hours (period of telophase or exit from mitosis) after drug release. We conclude that *bub1* is mitotically regulated and is highly expressed in mitosis.

EXAMPLE 8: Protein Localization of BUB1 Serine/Threonine Protein Kinase

Materials and Methods

The intracellular location of BUB1 protein was determined by immunoflourescence with BUB1-specific antibody. HeLa cells and H1299 cells were grown to confluence. Cells were trypsinized and then plated on 1 cm poly-lysine pre-treated glass cover slips in 60 mm plates and incubated at 37 °C for 1 day. Cells were washed with 5 mL ice-cold PBS after removing medium and then were fixed with -20 °C-cold 100% methanol. Methanol fixed cells can be used for a week or so. Cells were blocked with 30 μL immunoflourescence blocking buffer (3% BSA, 0.01% Triton X100 in PBS) for 30 min and washed with 0.01% Triton X100 in PBS. The purified anti-BUB1 peptide 506 antibody was diluted 1:100-fold in the blocking buffer. Rhodomine or fluorescein (FITC)-conjugated AffinPure F(ab)2 fragments of goat anti-rabbit IgG were

diluted 1:400-fold and used as secondary antibody. Dapi (62.5 µg/mL) was used for staining nuclei.

Results

Immunofluorescence experiments to determine the intracellular location of endogenous BUB1 were performed using the affinity purified anti-BUB1 antibody in both HeLa and H1299 cells. BUB1 protein exclusively localized to kinetochores during mitosis. BUB1 antibody stained all chromosome pairs during prophase and early metaphase in mitosis, demonstrating that BUB1 is a centromeric protein.

Immunofluorescence experiments were also performed in pcDNA3.1.BUB1 transiently transfected 293T cells. The results indicated that BUB1 protein does not co-localize with Crest, but rather appears to be bound to microtubules when the BUB1 protein is overexpressed following transient transfection. Cells overexpressing BUB1 are healthy and divide normally, but BUB1 protein does not localize to kinetochores.

EXAMPLE 9: BUB1 *in vitro* and *in vivo* Kinase Assays

The pcDNA3.1.BUB1.HA expression plasmid was transfected into human 293T cells. The crude protein lysates were extracted from the transfected cells and used for BUB1 kinase activity assays. Four different assays were performed: (1) BUB1 *in vitro* kinase autophosphorylation, (2) BUB1 phosphorylation of exogenous substrates, (3) BUB1 *in vivo* kinase activity, and (4) BUB1 phosphoamino acid analysis.

BUB1 *in vitro* Kinase Autophosphorylation Assay:

Materials and Methods

The pcDNA3.1 BUB1 expression plasmids (10 µg DNA/90 mm plate) containing HA-tagged wild-type (BUB1.HA), N-terminally truncated (BUB1.N.HA), or the kinase dead form of BUB1 (BUB1.KR.HA), were introduced into human 293T cells with lipofectamine (Gibco BRL, Gaithersburg, MD). After 48 hours, the cells were harvested in 0.4 mL of kinase lysis buffer (50 mM HEPES, pH 7.5, 100 mM KCl, 25 mM NaF, 1 mM sodium orthovanadate, 0.5 mM nonidet P-40 (Sigma, St. Louis, MO), 1 mM DTT, 170 µg/mL PMSF, 25 µg/mL leupeptin, 20 µg/mL trypsin inhibitor, 25 µg/mL aprotinin). Protein extracts were immunoprecipitated using anti-HA antibody (Babco, Berkeley, CA) at a final dilution of 1:750 and protein G beads (Boehringer Mannheim, Indianapolis, IN), then washed with kinase lysis buffer three times followed by two washes with kinase buffer (20 mM HEPES, pH 7.5, 150 mM KCl, 10 mM $MgCl_2$). Twenty µL kinase

reaction mix (1X kinase buffer, 1 mM DTT, 0.1 mM ATP, 10 μCi ^{32}P-r-ATP) was added to the washed beads and the mixture was incubated at 37 °C for 40 minutes with constant agitation. The reaction was stopped by adding 8 μL 5X PAGE buffer and boiled at 100 °C for 5 minutes.

Results

Kinase assays using BUB1.HA, BUB1.KR.HA (kinase dead), and BUB1.N.HA (dominant/negative) in 293T transient cell lysates were performed. The radio-labeled proteins were immunoprecipitated by anti-HA antibody and applied to a 10% polyacrylamide SDS gel. An 150 kDa protein was detected by autoradiography from the sample with the wild type BUB1 gene, but not in the samples of the kinase dead BUB1, the dominant/negative BUB1, or the vector only. The BUB1 protein contains kinase activity, and can be autophosphorylated.

BUB1 *in vitro* Phosphorylation of Exogenous Substrates

Materials and Methods

In vitro kinase assays using exogenous substrates were identical to the *in vitro* autophosphosylation assay, with the exception that the kinase reaction mix contained 1 mg/mL of histone H1, myelin basic protein, or alpha casein protein (Boehringer Mannheim). The N-terminal portion of BUB1 was also tested.

Results

The phosphorylation of histone H1 by BUB1 was strong, of myelin basic protein was weak, and no signal was detected for alpha casein by autoradiography using BUB1 which was obtained from the immunoprecipitated transiently transfected cell extracts. Histone H1 is a good *in vitro* substrate for the BUB1 protein kinase.

The *in vitro* kinase assay was also used to test whether the N-terminal, truncated form of BUB1 could also be a substrate for the BUB1 kinase, since a radioactively labeled BUB1.N protein (50 kDa) was detected in the *in vivo* kinase assay (herein), suggesting that BUB1.N protein is a substrate for an unidentified kinase or for the BUB1 kinase *in vivo*. Autoradiography of the assay using immunoprecipitated BUB1.N as the substrate for immunoprecipitated BUB1, showed that the BUB1.N protein could be detected (results not shown). Therefore, it appears that BUB1.N protein is a weak substrate for BUB1 kinase.

BUB1 *in vivo* Kinase Assay:

Materials and Methods

The pcDNA3.1.BUB1 expression plasmids (10 μg DNA/100 mM plate) containing HA-tagged full length, truncated, or the kinase dead form of BUB1, were introduced into human 293T cells with lipofectamine (Gibco BRL). After 48 hours, the cell medium was changed to phosphate-free medium (Gibco BRL) containing 10% dialyzed fetal calf serum (Gibco BRL). Twenty hours after the medium change, inorganic ^{32}P-phosphate at 2 mCi/100 mm plate was added to the medium and the cells were incubated for 3 hours. Twelve μL of pervanadate solution (50 mM sodium orthovanadate, 15% hydrogen peroxide) was added, and the cells were incubated for an additional 10 minutes. The cell culture dishes were placed on ice, and were rinsed with 4 °C phosphate-buffered saline. The crude protein lysates were extracted with 450 μL of kinase lysis buffer (50 mM HEPES, pH 7.5, 100 mM KCl, 25 mM NaF, 1 mM sodium orthovanadate, 0.5 mM nonidet P-40 (Sigma), 1 mM DTT, 170 μg/mL PMSF, 25 μg/mL leupeptin, 20 μg/mL trypsin inhibitor, 25 μg/mL aprotinin), and were immunoprecipitated using anti-HA antibody.

Results

The assay was performed by adding inorganic ^{32}P to 293T cells transiently transfected with BUB1 or BUB1.N. Crude protein was extracted from the radio-labeled cells. Cell lysates were immunoprecipitated by anti-HA antibody and then applied to a 10% polyacrylamide SDS gel. A phosphorylated band was observed at the position with molecular mass of 150 kDa corresponding to BUB1. No signal was detected in the vector only control. A weak signal was observed from the sample with BUB1.N transfection. This was not seen in the *in vitro* kinase assay.

BUB1 Kinase Phosphoamino Acid Analysis:

Material and Methods

Hunter's protocol was employed for BUB1 kinase phosphoamino acid analysis. Following the completion of an *in vitro* kinase assay, the radio-labeled BUB1 band was excised from the dried polyacrylamide gel. The gel slice was rehydrated in 0.5 mL of 50 mM ammonium bicarbonate, pH 7.3 to 7.6, for 5 minutes, and then was homogenized into a fine suspension. The volume of suspension was adjusted to 1 mL with 50 mM ammonium bicarbonate with a final concentration of 5% beta mercaptoethanol (v/v) and 0.1% SDS (w/v). The mixture was incubated at 100 °C for 5 minutes followed by 37 °C for 12 hours with agitation. The particulate fraction was collected by centrifugation and re-extracted with 50 mM ammonium bicarbonate, pH 7.3 to 7.6, 5% beta mercaptoethanol

(v/v), 0.1% SDS (w/v), by incubation at 100 °C for 5 minutes followed by 37 °C for 60 minutes. Following centrifugation, the liquid extract was combined with the prior extract to give a total volume of 1.2 mL. The extract was centrifuged again to remove particulate material, and the supernatant was isolated and chilled to 4 °C. A total of 20 µg of bovine serum albumin was added to the extract and the proteins were precipitated by addition of 250 µL of 100% trichloroacetic acid and incubated at 4 °C for 60 minutes. The precipitated proteins were isolated by centrifugation, and the resulting pellets were washed with 500 µL of 4 °C 100% acetone. The protein pellets were air dried, and then were dissolved in 100 µL of 5.7 M HCl. The solution was heated at 100 °C for 60 minutes in a screw-capped tube, and then was dried under vacuum. The residual pellet was dissolved in formic acid (2% v/v) and glacial acetic acid (7.3% v/v) containing 67 µg/mL phosphoserine, 67 µg/mL phosphothreonine, 67 µg/mL phosphotyrosine, pH 1.9. A sample containing a total of 600 cpm was spotted onto a cellulose TLC plate (VWR scientific, San Francisco, CA). Two dimension chromatography was performed. The first dimension gel was run in a buffer of formic acid (2% v/v) and glacial acetic acid (7.3% v/v), pH 1.9, for 25 minutes at 1.0 kV. The plate was air-dried, then subjected to the second dimension gel electrophoresis in a buffer of glacial acetic acid (5% v/v), pyridine (0.5% v/v), 0.5 mM EDTA, pH 3.5 for 16 minutes at 1.3 kV. The plate was air-dried again, then sprayed with 0.25% ninhydrin in acetone and developed for 15 minutes at 65 °C. The position and identity of the labeled amino acids were determined by autoradiography.

Results

The Hunter's protocol was performed to determine the phosphorylation sites (amino acid) for the BUB1 protein kinase. Radio-labeled amino acids were observed at the locations of phosphoserine and phosphothreonine, but no signal was seen at the position of phosphotyrosine. Therefore, BUB1 kinase is a serine/threonine protein kinase.

EXAMPLE 10: Development of High Through-Put Screening Protocol for BUB1 Protein Kinase Inhibitors

Since BUB1 RNA is overexpressed in many cancer cells, presumably BUB1 kinase activity is upregulated in many cancer cells, causing a loss of the mitotic check point. Inhibition of BUB1 kinase activity might be a way to inhibit uncontrolled cell

division and therefore to block tumor cell growth. Overexpression of recombinant BUB1 protein can be employed for *in vitro* kinase inhibitor screening.

Recombinant baculoviruses have become widely used as vectors to express heterologous genes in cultured insect cells. Heterologous genes placed under the control of the strong polyhedrin promotor of the *Autograhpa california* nuclear polyhidrosis virus (AcNPV) are often abundantly expressed during the late stages of infection. In most cases, the recombinant proteins are processed, modified, and targeted to their appropriate cellular locations, where thy are functionally similar to their authentic counterparts.

Expression of Recombinant BUB1 in Baculovirus System

Materials and Methods

The baculovirus expression system was purchased from Gibco BRL, Gaithersburg, MD. BUB1.HA wild type and BUB1.KR.HA (kinase dead) were subcloned into pFastBacHtb (with the histidine tag at the N-terminal) and pFastBacG2T (the glutathione transferase gene replaced the histidine tag). The recombinant plasmids were constructed in the regular *E. coli* host strain, XL2-blue. The bacmids were then propagated in *E. coli* DH10Bac cells. The recombinant bacmid DNA was isolated and transfected into insect cells. Recombinant baculovirus particles were then used to infect insect cells to overexpress recombinant proteins.

Results

72 hours after baculovirus infection, insect cells were harvested and crude protein lysates were extracted. Expression of recombinant BUB1 proteins was determined by Western analysis using anti-HA antibody. The expected BUB1 bands were detected with molecular mass 150 kDa. Both recombinant proteins of BUB1 wild type, and kinase dead fused to HISx6 and GST, were expressed in the baculovirus system.

Baculoviral Expressed BUB1 *in vitro* Kinase Activity Assay

Materials and Methods

Protein extracts from overexpressed recombinant BUB1 protein insect cells were used for BUB1 *in vitro* kinase activity assays. The protocol for the kinase assay was the same as described in Example 9.

Results

Protein extracts from BUB1 and BUB1.KR fused to HIS tags were isolated from insect cells, and the *in vitro* kinase activity assay was performed. The radioactively labeled protein was immunoprecipitated by anti-HA antibody and then applied onto a 10%

polyacrylamide SDS gel. The 150 kDa BUB1 protein was observed only in the sample containing recombinant wild-type BUB1. The recombinant BUB1 can be autophosphorylated and can phosphorylate the exogenous substrate histone H1.

BUB1 overexpressed in the baculovirus system provides a powerful tool for a high throughput screen. Small chemical molecules are one example of compounds to be screened for their use as therapeutic drugs.

One skilled in the art would readily appreciate that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. The molecular complexes and the methods, procedures, treatments, molecules, specific compounds described herein are presently representative of preferred embodiments are exemplary and are not intended as limitations on the scope of the invention. Changes therein and other uses will occur to those skilled in the art which are encompassed within the spirit of the invention are defined by the scope of the claims.

It will be readily apparent to one skilled in the art that varying substitutions and modifications may be made to the invention disclosed herein without departing from the scope and spirit of the invention.

All patents and publications mentioned in the specification are indicative of the levels of those skilled in the art to which the invention pertains.

The invention illustratively described herein suitably may be practiced in the absence of any element or elements, limitation or limitations which is not specifically disclosed herein. Thus, for example, in each instance herein any of the terms "comprising", "consisting essentially of" and "consisting of" may be replaced with either of the other two terms. The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed.

In particular, although some formulations described herein have been identified by the excipients added to the formulations, the invention is meant to also cover the final formulation formed by the combination of these excipients. Specifically, the invention includes formulations in which one to all of the added excipients undergo a reaction during formulation and are no longer present in the final formulation, or are present in modified forms.

In addition, where features or aspects of the invention are described in terms of Markush groups, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group. For example, if X is described as selected from the group consisting of bromine, chlorine, and iodine, claims for X being bromine and claims for X being bromine and chlorine are fully described.

EXAMPLE 11: Antibody characterization:

Materials and Methods

A rabbit polyclonal antibody was raised against a specific peptide sequence within the BUB1 coding region. The peptide (506A) consists of the following amino acid sequence corresponding to residues 385 to 498 of the BUB1 amino acid sequence: DDKDEWQSLDQNED (SEQ ID NO:19). The IgG fraction was purified by affinity chromatography using protein A agarose beads by standard methodologies. The IgG fraction was dialyzed against PBS and adjusted to 1 mg/ml BSA and frozen in aliquots at –20 °C.

Results

Specificity of the antibodies was confirmed using recombinant BUB1 protein expressed in human 293T cells. The antibodies detect a protein of approximately 150 kDa protein both by immunoprecipitation and immunobloting experiments. The signal was competed by including the immunizing peptide with the antibody. The BUB1 antibody immunoprecipitated a kinase activity that could transphosphorylate histone H1. BUB1 was also autophosphorylated in kinase assays of BUB1 immunocomplexes.

EXAMPLE 12: Cell cycle regulation.

Materials and Methods

HeLa cells were synchronized by a double thymidine/aphidicolin block at the G1/S transition. The cells were released from the block and time points were taken as the cells proceed through the cell division cycle.

Results

The amount of BUB1 protein as well as the associated kinase activity is cell cycle regulated peaking in mitotis. The activity appears to be maximal during metaphase corresponding to maximal activation of $p34^{cdc2}$ (CDK1). Consistent with this observation

the activity is also induced by nocodozole treatment which arrest the cells in G2 and mitosis, but not by serum starvation or treatment with thymidine which induce a G1 arrest.

EXAMPLE13: Antisense Oligonucleotides

Materials and Methods

An antisense oligonucleotide was identified that specifically down regulates BUB1 protein levels. The sequence of this oligo is: GTGCCAAGCCAGCAGATAAAT (SEQ ID NO:24). This corresponds to nucleotides 2889 to 2909 of the hBUB1 cDNA sequence.

Results

This oligonucleotide specifically down regulates BUB1 protein levels as determined by immunobloting with BUB1 specific antisera. The level of β-tubulin remains unchanged. Other mitotic proteins such as Aurora1, Aurora2, cyclin A, cyclin B, and polo-like kinase 1 (PLK1) were up-regulated in cells treated with the BUB1 antisense oligonucleotide as compared to control cells. This is consistent with a mitotic arrest of the cells. To confirm this the DNA content of the BUB1 antisense oligonucleotide-treated cells was examined by flow cytometry. The analysis showed that the BUB1 antisense oligonucleotide-treated cells were arrested with a 4N DNA content consistent with a mitotic arrest. To further characterize the arrest point, treated cells were stained with DAPI to stain the DNA as wells with antibodies to α-tubulin and BUB1 to stain microtubules and the BUB1 protein, respectively. Cells that did not express BUB1 appeared to arrest in mitosis, specifically in metaphase with condensed chromosomes aligned on the metaphase plate and a normal bipolar mitotic spindle. In a small percentage of these cells the chromosome appeared to randomly disperse as if the chromosomes were no longer attached to the mitotic spindle. This is consistent with BUB1 protein being involved in microtubule binding to the kinetocore of the chromosome. The cell number of the BUB1 antisense-treated cells decreased significantly two days post-treatment implying that loss of BUB1 was lethal. To examine whether this loss of cell number was due to apoptosis, a TUNEL assay was performed on the cells treated with the BUB1 antisense oligonucleotide. The TUNEL assay measures the amount of DNA fragmentation of a given cell population which is a hallmark of apoptotic cell death. The cells treated with control antisense oligonucleotide had no increase in DNA fragmentation for up to three days post-treatment, whereas the BUB1 antisense oligonucleotide-treated cells had a 10-

fold increase in apoptosis two days post-treatment. These observations support that BUB1 in an essential gene required for mitosis.

EXAMPLE 14: BUB1 Cooperates with Ras.

Materials and Methods

To examine the effects of overexpression of BUB1 on cellular transformation mouse NIH3T3 cells were transfected with activated H-ras (valine12) alone or in combination with wild-type BUB1 (BUB1WT) or a N-terminal deletion mutant corresponding to amino acids 1-332 (BUB1N).

Results

In three independent experiments, the combination of activated H-ras and BUB1WT generated twice as many foci as activated H-ras alone or activated H-ras plus BUB1N. These data suggest that BUB1 can cooperate with the ras oncogene to transform cells.

EXAMPLE 15: Isolation of cDNAs Encoding Mammalian NEK-Related Protein Kinases

Materials And Methods

Makegene Bioinformatics EST Assembler

The EST reports were downloaded from ncbi (www.ncbi.nlm.nih.gov). After decompressing the files, the program 'report2est' was used to extract the following information: 1) EST names, 2) GenBank Accession numbers, 3) GenBank gi numbers, 4) Clone Id numbers, 5) the nucleotide sequences of the ESTs, 6) the organism, 7) the library name, 8) the name of the lab, and 9) the institution. The output of 'report2est' is a file in FASTA format with all of the information listed above in the first line of each entry except the sequence, which is listed in the second line of each entry. The resulting file was formatted for BLAST using 'pressdb' (available as part of the ncbi tool kit).

To build a gene or part of a gene from EST's, the program 'Makegene' was used. Input to this program is a query sequence and the organism/species for which a gene is to be built. An initial search of the formatted EST database described above was performed using BLAST (blastn). Any results that contained warnings, such as polyA tails or other repeat elements, were eliminated from future queries. The program 'blast_parse_reports' was used to extract the FASTA header line from the search results and the output was then filtered to extract only FASTA header lines for the desired species.

The initial results, having been filtered for warnings and species, went into a loop in which searches against the database are repeated until no new EST's are found. The loop consists of the following steps: 1) when possible, the names of both ends of the EST's are extracted from the database by searching using the 'Clone Id' field or the part of the 'EST name' field before the .r or .s postscript, 2) any EST's that have been used as queries in previous loops are removed from the current query by the program 'subtract', 3) the resulting list of EST's is used to extract the sequences from the database by the program batch_parse_fasta, 4) BLAST is run against the database using each sequence, 5) the output files from BLAST containing warnings are removed, 6) the results are filtered by species, and 7) the loop is reentered if there were new EST's found in the last pass through the loop.

The EST's chosen by 'Makegene' were used as input for the program 'mpd2_cluster' (Hide, W., Burke, J, and Davison, D.U. of Houston, unpublished) which clusters overlapping sequences. The programs 'contig' (Kerlavage, T., TIGR, unpublished), 'gde2mult', and 'gde2sing' [Smith, S.W., *et al.*, CABIOS 10:671-675 (1994)], were used to make an alignment and consensus sequence of the overlapping EST's.

DNA Sequencing

Plasmid DNA was purified using Qiagen columns, and was sequenced using cycle-sequencing energy transfer BIG DYE-terminator chemistry with Taq, FS DNA polymerase (ABI, Foster City, CA). Sequencing reactions were run on an ABI Prism 377 DNA sequencer, and were analyzed using the BLAST alignment algorithm (Altschul, S.F. *et al.* (1990) J. Mol. Biol. 215:403-410). Primer-walking sequencing oligonucleotides were synthesized on an Applied Biosystems 3948 DNA synthesizer using established phosphoramidate chemistry (ABI, Foster City, CA).

Results

cDNA Cloning and Characterization of NEK4a

The human *nek4a* cDNA sequence is represented by the EST fragment AA214523, identified through a Smith-Waterman search of the EST database. The complete sequence of this clone was determined, and was used to generate the predicted full-length NEK4a amino acid sequence (SEQ ID. NO:8)

The full-length human *nek4a* cDNA (SEQ ID NO:7) is 1,449 bp long, and consists of a 1,107 bp open reading frame (ORF) flanked by a 121 bp 5' UTR and a 200 bp 3' UTR

that is followed by a 21 nucleotide polyadenylated region. No polyadenylation signal (AATAAA) was found preceding the polyadenylated region. The sequence flanking the first ATG conforms to the Kozak consensus (Kozak, M. (1987) Nucleic Acids Res. 15:8125-8148) for an initiating methionine, and is believed to be the translational start site for NEK4a.

Several EST fragments span the complete *nek4a* sequence with AA214523 at the 5' end and AA147800 at the 3' end.

cDNA Cloning and Characterization of NEK4b

The human *nek4b* cDNA sequence is represented by the EST fragment AA147800, identified through a Smith-Waterman search of the EST database. The complete sequence of this clone was determined, and was used to generate the predicted full-length NEK4b amino acid sequence (SEQ ID. NO:10).

The partial human *nek4b* cDNA (SEQ ID NO:9) is 1,346 bp long, and consists of a 804 bp ORF flanked by 499 bp 3' UTR that is followed by a 41 nucleotide polyadenylated region. No polyadenylation signal (AATAAA) was found preceding the polyadenylated region. Since the coding region is open throughout the 5' extent of this sequence, this appears to be a partial cDNA lacking the N-terminal start methionine.

Several EST fragments span the *nek4b* sequence, with N30324 at the 5' end and AA461376 at the 3' end.

cDNA Cloning and Characterization of NEK5

The EST fragment, AA181165, which represents the entire human *nek5* gene, was identified using a Smith-Waterman search of the EST. The complete sequence of the clone was determined, and was used to generate the predicted full-length NEK5 amino acid sequence (SEQ ID. NO:4).

The full-length human nek5 cDNA (SEQ ID NO:3) is 3,854 bp long. The human NEK5 cDNA contains a 906 bp ORF flanked by a 31 bp 5' UTR (1-31) and a 2,894 bp 3' UTR (941-3,834) that is followed by a 20 nucleotide polyadenylated region. A polyadenylation signal (AATAAA) is found at positions 3817-3822. The sequence flanking the first ATG conforms to the Kozak consensus for an initiating methionine, and is believed to be the translational start site for NEK5.

At least one EST fragment, AA181165, spans the complete nek5 sequence. AA181165 contains a 3,854 bp insert and an ORF of 906 bp (302 amino acids).

cDNA Cloning and Characterization of NEK6

The EST fragment, AA030068, which represents the partial human NEK6 gene was identified using a Smith-Waterman search of the EST database. The complete sequence of the clone was determined, and was used to generate the partial human NEK6 sequence.

The partial murine *nek6* cDNA (SEQ ID NO:5) is 1,310 bp long. The murine NEK6 cDNA contains a 907 bp ORF preceded by a 403 bp 5' UTR (1-403). The sequence flanking the first ATG weakly conforms to the Kozak consensus for an initiating methionine However, it is believed to be the translational start site for NEK6, for four reasons: (1) there is an adenine at the -3 position prior to the start codon, although no pyrimidine at the +4 position; (2) there are multiple in-frame and out-of-frame stop codons 5' of the putative start site; (3) there are no other possible start codons in any frame 5' of the putative strat codon; and (4) the sequence preceding the start site is 57.57% Gs and Cs, moderately characteristic of a 5'UTR region (highly GC- rich).

One EST fragment, AA030068, represents the partial *nek6* sequence. EST clone AA030068 contains a 1,310 bp insert and an ORF of 907 bp (302 amino acids).

EXAMPLE 16: Expression Analysis Of Mammalian Nek-Related Protein Kinases

Materials And Methods

Dna Array-Based Expression Analysis

Plasmid DNA array blots were prepared by loading 0.5 µg denatured plasmid for each NEK-related kinase on a nylon membrane. The [α^{32}P]dCTP-labeled single-stranded DNA probes were synthesized from total RNA isolated from several human and rodent immune sources, including primary thymus (B6Thy), mast (MC/9), and T cells (CD4+), as well as from the following cell lines: T (MOLT-4, CEM/C1, Jurkat), B (2PK3, WEHI-231, DAUDI, RAMOS, RAJI), TH2 (D10.G4.1, pL104, pL3), TH1 (pGL10), CTL (L3), and basophil (RBL-2H3). Hybridization was performed at 42 °C for 16 hours in 6X SSC, 0.1% SDS, 1X Denhardt's solution, 100 µg/mL denatured herring sperm DNA with 10^6 cpm/mL of [α^{32}P]dCTP labeled single-stranded probe. The filters were washed in 0.1X SSC/0.1% SDS, 65 °C, and exposed for quantitative analysis on a Molecular Dynamics phosphorimager.

RESULTS

Distribution of NEK-related Gene Transcripts in an Immune Cell RNA Panel

Expression of the three NEK-related kinases was examined in a panel of human and rodent immune cell sources by hybridization to a DNA-array blot containing plasmids encoding each of these genes (NEK4 EST GB:H59937, NEK5 EST GB:AA102912, and NEK6 EST GB:AA030253). (The NEK4 probe used in this experiment should detect the NEK4a and NEK4b isoforms with equal efficiency).

NEK4 was broadly expressed in the entire immune cell panel with highest and lowest expression observed in mast cells and thymus, respectively. NEK5 was broadly expressed in the entire immune cell panel with highest and lowest expression observed in TH1 cells and thymus, respectively. NEK6 was broadly expressed in the entire immune cell panel with highest and lowest expression observed in TH1 and basophil cells, respectively.

This analysis suggests that the novel NEK-subfamily kinases may be candidate targets for various immune disorders. Their broad expression pattern also suggests that they may mediate functions vital to the basic biology of most proliferating cells of the immune system.

Table 2
Expression of Novel NEK-subfamily Kinases
in an Immune Cell Panel

Immune Source	Cell Type	NEK4	NEK5	NEK6
		Log CPM		
m B6Thy	Thymus	2.56	3.49	3.68
r basophils	Basophil	3.83	4.86	2.22
m MC/9	Mast	4.06	5.19	4.87
h mast cells	Mast	5.37	4.3	4.87
m 2PK3	B	4.04	5.33	5.17
m WEHI-231	B	3.5	5.23	4.81
h DAUDI	B	4.43	4.15	4.43
h RAMOS	B	4.15	4.04	4.26
h RAJI	B	4.72	4.24	4.5
m D10.G4.1	TH2	4.18	5.35	5.5
mL3	CTL	4.7	4.38	4.47
m pL1O4	TH2	4.22	5.19	5.07
m pL3	TH2	3.82	5.88	4.61
m pGL10	TH1	4.77	6.34	5.49
h MOLT-4	T	4.04	3.93	4.31
h CEM/C1	T	3.65	3.96	3.99
h JURKAT	T	3.8	4.08	4.34
h CD4+	T	4.09	3.79	3.51

EXAMPLE 17: *nek* Gene Expression Plasmid: Construction and Recombinant NEK Expression

Expression Plasmid Construction

Expression constructs can be generated for NEK kinases including, for example: a) pcDNA3.1 expression plasmids containing: full-length *nek* nucleotide base sequences with and without an HA tag; the nucleotide base sequence encoding the N-terminal domains (amino acid 1 to 42, 1-31, or 1-22) of NEK4a, Nek5, or NEK6, respectively, with and without an HA tag; the nucleotide base sequence encoding the catalytic domains (amino acid 43-294, 1-259,32-283, or 23-302) of NEK4a, NEK4b, NEK5, or NEK6, respectively, with and an without HA tag, and the C-terminal domains (amino acids 295-313, 260-268, or 284-302) of NEK4a, NEK4b, or NEK5, respectively, with and without an HA tag; b) pGEX 4T-3 fusion plasmids containing: the full length nucleotide base sequence of *nek*; the nucleotide base sequence encoding the N-terminal domain the nucleotide base sequence encoding the N-terminal domains (amino acid 1 to 42, 1-31, or 1-22) of NEK4a, Nek5, or NEK6, respectively, the nucleotide base sequence encoding the catalytic domains (amino acid 43-294, 1-259,32-283, or 23-302) of NEK4a, NEK4b,

NEK5, or NEK6, respectively, and the C-terminal domains (amino acids 295-313, 260-268, or 284-302) of NEK4a, NEK4b, or NEK5, respectively; and c) pEGFP fusion plasmid containing the full length nucleotide base sequence encoding NEK kinases. The NEK N-terminal domains are predicted to function as dominant negative constructs, and can be used to further elucidate the function of NEK. The full-length genes can easily be obtained by methods well-known in the art given the partial sequences available.

EXAMPLE 18: Recombinant Expression of NEK Protein Kinases

Materials and Methods

Transient Expression of NEK in Mammalian Cells

The pcDNA expression plasmids (10 μg DNA/100 mm plate) containing sequences encoding the full length or truncated form of a NEK kinase, are introduced into human 293 cells with lipofectamine (Gibco BRL). After 72 hours, the cells are harvested in 0.5 mL solubilization buffer (20 mM HEPES pH 7.4, 150 mM NaCl, 10% glycerol, 1% Triton X-100, 1 mM EGTA, 2 mM phenylmethylsulfonyl fluoride, 1 μg/mL aprotinin). Sample aliquots are resolved by SDS polyacrylamide gel electrophoresis (PAGE) on 7.5% acrylamide/0.5% bis-acrylamide gels and are electrophoretically transferred to nitrocellulose. Non-specific binding is blocked by pre-incubating blots in Blotto (phosphate buffered saline containing 5% w/v non-fat dried milk and 0.2% v/v nonidet P-40 (Sigma)). Recombinant protein is detected using various NEK antisera, or anti-HA specific antisera, by Western analysis. NEK-antisera is diluted approx. 1:1,000-fold, and anti-HA monoclonal antibody (Babco, Berkeley, California) is diluted 1:10,000-fold for the primary antibody reactions of Western blotting. Goat anti-rabbit HRP (BioRad, Richmond, California) is used as a secondary antibody for the NEK-antisera, and goat anti-mouse HRP (BioRad, Richmond, California) was used as a secondary antibody for the anti-HA antibody. ECL reagent (Amersham Life Science,) is used for detection.

Cell lysates (30 μg each) are applied to 7.5% SDS PAGE followed by Western blotting.

EXAMPLE 19: Generation of Specific Immunoreagents to NEK Protein Kinases

Materials and Methods

Specific immunoreagents are raised in rabbits against MAP-conjugated synthetic peptides corresponding to NEK kinases. Each peptide is injected into two rabbits.

Additional immunoreagents are generated by immunizing rabbits with the bacterially expressed GST-fusion proteins containing the N-terminal or C-terminal domain of NEK kinases.

Affinity Purification Of Anti-NEK Peptide Antibody

Ten mg NEK peptide is covalently coupled to 2 mL Immunolink (Pierce AminoLink Plus). The link consists of an aldehyde support arm that forms a covalent bond to primary amines via a Schiff base intermediate. The NEK peptide (ligand) is immobilized to AminoLink Plus coupling gel. The remaining active sites are blocked by adding the reducing agents. The column is then washed prior to use.

Twenty mL of filtered anti-NEK antiserum is applied to the immobilized antigen column. The purified anti-NEK IgG is eluted with 1 M propionic acid, is neutralized with 3 M Tris, pH 7.4 while eluting, and is buffered with 100 mM TEA buffered with 1 M Tris pH 6.8. The purified antibody is dialyzed against phosphate buffered saline pH 7.4, and is stored in 20% glycerol and 0.03% NaN_3 in PBS.

Anti-NEK antisera is tested by Western analysis using GST-NEK fusion proteins.

An affinity purified anti-NEK antibody is tested by Western analysis using a GST-NEK fusion protein, or mammalian cell lysates from transiently transfected and expressed NEK protein.

EXAMPLE 20: Endogenous Expression of NEK Protein Kinase

Materials and Methods

Multiple human tumor cell lines (HT-29, Calu-6, HCT15, MCF7, A549, C33a, HeLa, HT-29, H1299, SH422C, COLO205, SW480, SW837, SW948, and SNB19, for example) are grown to confluence. Protein lysates are isolated in HNTG buffer (20 mM HEPES, pH 7.4, 150 mM NaCl, 10% glycerol, 1% Triton X-100, 1 mM EGTA, 2 mM phenylmethylsulfonylfluoride, 1 μg/mL aprotinin). The protein concentration is determined by BioRad assay (BioRad, Richmond, California). Protein from each sample (30 μg each) is tested by Western analysis with the purified anti-NEK antibody.

EXAMPLE 21: Cell Cycle Regulation of NEK

Materials and Methods

HeLa cells are synchronized by adding 0.5 M thymidine and 250 μg/mL aphidicolin. Cells are harvested at 0, 2, 4, 6, 8, 9, 10, 11, and 12 hours after thymidine and aphidicolin are removed. Other experiments are performed by blocking HeLa cells at different cell cycle stages by serum starvation, with thymidine (G1/S/G2), or with nocodozole (M). Total RNA is isolated from the synchronized cells and cell cycle blocked cells. Total RNA (10 μg) is loaded on the nylon membranes. The RNA blots are hybridized with a *nek* probe.

Northern blots containing RNA derived from HeLa cells either in an exponential growth state or arrested in various stages of the cell cycle are probed for expression of *nek*.

EXAMPLE 22: Protein Localization of NEK Protein Kinase

Materials and Methods

The intracellular location of NEK protein is determined by immunoflourescence with NEK-specific antibody. HeLa cells and H1299 cells are grown to confluence. Cells are trypsinized and are then plated on 1 cm poly-lysine pre-treated glass cover slips in 60 mm plates and are incubated at 37 °C for 1 day. Cells are washed with 5 mL ice-cold PBS after removing medium and then are fixed with -20 °C-cold 100% methanol. Methanol fixed cells can be used for a week or so.

Cells are blocked with 30 μL immunoflourescence blocking buffer (3% BSA, 0.01% Triton X100 in PBS) for 30 min and are washed with 0.01% Triton X100 in PBS. The purified anti-NEK antibody is diluted approx. 1:100-fold in the blocking buffer. Rhodomine or fluorescein (FITC)-conjugated AffinPure F(ab)2 fragments of goat anti-rabbit IgG are diluted 1:400-fold and are used as secondary antibody. Dapi (62.5 μg/mL) is used for staining nuclei.

Immunofluorescence experiments to determine the intracellular location of endogenous NEK are performed using the affinity purified anti-NEK antibody in both HeLa and H1299 cells, for example. Immunofluorescence experiments are also performed in pcDNA3.1.NEK transiently transfected 293T cells.

EXAMPLE 23: NEK *in vitro* and *in vivo* Kinase Assays

The pcDNA3.1.NEK.HA expression plasmids are transfected into human 293T cells. The crude protein lysates are extracted from the transfected cells and are used for

NEK kinase activity assays. Four different assays are performed: (1) NEK *in vitro* kinase autophosphorylation, (2) NEK phosphorylation of exogenous substrates, (3) NEK *in vivo* kinase activity, and (4) NEK phosphoamino acid analysis.

NEK *in vitro* Kinase Autophosphorylation Assay:

Materials and Methods

The pcDNA3.1 NEK expression plasmids (10 μg DNA/90 mm plate), containing HA-tagged wild-type (NEK.HA) or N-terminally truncated (NEK.N.HA), are introduced into human 293T cells with lipofectamine (Gibco BRL, Gaithersburg, MD). After 48 hours, the cells are harvested in 0.4 mL of kinase lysis buffer (50 mM HEPES, pH 7.5, 100 mM KCl, 25 mM NaF, 1 mM sodium orthovanadate, 0.5 mM nonidet P-40 (Sigma, St. Louis, MO), 1 mM DTT, 170 μg/mL PMSF, 25 μg/mL leupeptin, 20 μg/mL trypsin inhibitor, 25 μg/mL aprotinin). Protein extracts are immunoprecipitated using anti-HA antibody (Babco, Berkeley, CA) at a final dilution of 1:750 and protein G beads (Boehringer Mannheim, Indianapolis, IN), then are washed with kinase lysis buffer three times, followed by two washes with kinase buffer (20 mM HEPES, pH 7.5, 150 mM KCl, 10 mM $MgCl_2$). Twenty μL kinase reaction mix (1X kinase buffer, 1 mM DTT, 0.1 mM ATP, 10 μCi ^{32}P-r-ATP) is added to the washed beads, and the mixture is incubated at 37 °C for 40 minutes with constant agitation. The reaction is stopped by adding 8 μL 5X PAGE buffer, and is boiled at 100 °C for 5 minutes.

Kinase assays using NEK.HA and NEK.N.HA (dominant/negative) in 293T transient cell lysates are performed. The radio-labeled proteins are immunoprecipitated by anti-HA antibody, and are applied to a 10% polyacrylamide SDS gel.

NEK *in vitro* Phosphorylation of Exogenous Substrates

Materials and Methods

In vitro kinase assays using exogenous substrates are identical to the *in vitro* autophosphosylation assay, with the exception that the kinase reaction mix contains 1 mg/mL of histone H1, myelin basic protein, or alpha casein protein (Boehringer Mannheim).

NEK *in vivo* Kinase Assay:

Materials and Methods

The pcDNA3.1.NEK expression plasmids (10 μg DNA/100 mM plate), containing HA-tagged full length or truncated NEK, are introduced into human 293T cells with lipofectamine (Gibco BRL). After 48 hours, the cell medium is changed to phosphate-free

medium (Gibco BRL), containing 10% dialyzed fetal calf serum (Gibco BRL). Twenty hours after the medium change, inorganic ^{32}P-phosphate at 2 mCi/100 mm plate is added to the medium, and the cells are incubated for 3 hours. Twelve µL of pervanadate solution (50 mM sodium orthovanadate, 15% hydrogen peroxide) is added, and the cells are incubated for an additional 10 minutes. The cell culture dishes are placed on ice, and are rinsed with 4 °C phosphate-buffered saline. The crude protein lysates are extracted with 450 µL of kinase lysis buffer (50 mM HEPES, pH 7.5, 100 mM KCl, 25 mM NaF, 1 mM sodium orthovanadate, 0.5 mM nonidet P-40 (Sigma), 1 mM DTT, 170 µg/mL PMSF, 25 µg/mL leupeptin, 20 µg/mL trypsin inhibitor, 25 µg/mL aprotinin), and are immunoprecipitated using anti-HA antibody.

The assay is performed by adding inorganic ^{32}P to 293T cells transiently transfected with NEK or NEK.N. Crude protein is extracted from the radio-labeled cells. Cell lysates are immunoprecipitated by anti-HA antibody, and are then applied to a 10% polyacrylamide SDS gel.

NEK Kinase Phosphoamino Acid Analysis:

Material and Methods

Hunter's protocol is employed for NEK kinase phosphoamino acid analysis. Following the completion of an *in vitro* kinase assay, the radio-labeled NEK band is excised from the dried polyacrylamide gel. The gel slice is rehydrated in 0.5 mL of 50 mM ammonium bicarbonate, pH 7.3 to 7.6, for 5 minutes, and then is homogenized into a fine suspension. The volume of suspension is adjusted to 1 mL with 50 mM ammonium bicarbonate with a final concentration of 5% beta mercaptoethanol (v/v) and 0.1% SDS (w/v). The mixture is incubated at 100 °C for 5 minutes followed by 37°C for 12 hours with agitation. The particulate fraction is collected by centrifugation and re-extracted with 50 mM ammonium bicarbonate, pH 7.3 to 7.6, 5% beta mercaptoethanol (v/v), 0.1% SDS (w/v), by incubation at 100 °C for 5 minutes followed by 37 °C for 60 minutes. Following centrifugation, the liquid extract is combined with the prior extract to give a total volume of 1.2 mL. The extract is centrifuged again to remove particulate material, and the supernatant is isolated and chilled to 4 °C. A total of 20 µg of bovine serum albumin is added to the extract and the proteins are precipitated by addition of 250 µL of 100% trichloroacetic acid and incubated at 4 °C for 60 minutes. The precipitated proteins are isolated by centrifugation, and the resulting pellets are washed with 500 µL of 4 °C 100%

acetone. The protein pellets are air dried, and then are dissolved in 100 µL of 5.7 M HCl. The solution is heated at 100 °C for 60 minutes in a screw-capped tube, and then is dried under vacuum. The residual pellet is dissolved in formic acid (2% v/v) and glacial acetic acid (7.3% v/v) containing 67 µg/mL phosphoserine, 67 µg/mL phosphothreonine, 67 µg/mL phosphotyrosine, pH 1.9. A sample containing a total of approx. 600 cpm is spotted onto a cellulose TLC plate (VWR scientific, San Francisco, CA). Two dimension chromatography is performed. The first dimension gel is run in a buffer of formic acid (2% v/v) and glacial acetic acid (7.3% v/v), pH 1.9, for 25 minutes at 1.0 kV. The plate is air-dried, then subjected to the second dimension gel electrophoresis in a buffer of glacial acetic acid (5% v/v), pyridine (0.5% v/v), 0.5 mM EDTA, pH 3.5 for 16 minutes at 1.3 kV. The plate is air-dried again, then sprayed with 0.25% ninhydrin in acetone and is developed for 15 minutes at 65 °C. The position and identity of the labeled amino acids are determined by autoradiography.

The Hunter's protocol is performed to determine the phosphorylation sites (amino acid) for the NEK protein kinase.

EXAMPLE 24: Development of High Through-Put Screening Protocol for NEK Protein Kinase Inhibitors

Since NEK is involved in cell-cycle regulation, deregulation of these kinases through overexpression and/or mutational events plays a role in cancer, for example. Inhibition of NEK kinase activity will inhibit uncontrolled cell division and thereby block tumor cell growth. Overexpression of recombinant NEK protein can be employed for *in vitro* kinase inhibitor screening.

Recombinant baculoviruses have become widely used as vectors to express heterologous genes in cultured insect cells. Heterologous genes placed under the control of the strong polyhedrin promotor of the *Autograhpa california* nuclear polyhidrosis virus (AcNPV) are often abundantly expressed during the late stages of infection. In most cases, the recombinant proteins are processed, modified, and targeted to their appropriate cellular locations, where they are functionally similar to their authentic counterparts.

Expression of Recombinant NEK in a Baculovirus System

Materials and Methods

The baculovirus expression system is purchased from Gibco BRL, Gaithersburg, MD. NEK.HA wild type and NEK.KR.HA (a kinase dead version) are subcloned into

pFastBacHtb (with the histidine tag at the N-terminal) and pFastBacG2T (the glutathione transferase gene replaces the histidine tag). The recombinant plasmids are constructed in the regular *E. coli* host strain, XL2-blue. The bacmids are then propagated in *E. coli* DH10Bac cells. The recombinant bacmid DNA is isolated and transfected into insect cells. Recombinant baculovirus particles are then used to infect insect cells to overexpress recombinant proteins.

Kinase dead constructs can be made by methods well-known in the art, including point mutations, insertion and deletion of nucleic acids such that kinase activity is lost.

72 hours after baculovirus infection, insect cells are harvested and crude protein lysates are extracted. Expression of recombinant NEK proteins is determined by Western analysis using anti-HA antibody.

Baculoviral Expressed NEK *in vitro* Kinase Activity Assay

Materials and Methods

Protein extracts from overexpressed recombinant NEK protein insect cells were used for NEK *in vitro* kinase activity assays. The protocol for the kinase assay was the same as described in Example 9.

Protein extracts from NEK and NEK.KR fused to HIS tags are isolated from insect cells, and the *in vitro* kinase activity assay is performed. The radioactively labeled protein is immunoprecipitated by anti-HA antibody and then is applied onto a 10% polyacrylamide SDS gel.

NEK overexpressed in the baculovirus system provides a powerful tool for a high throughput screen. Small chemical molecules are one example of substances to be screened for their use as therapeutic drugs.

EXAMPLE 25: NEK Antisense Oligos

Antisense oligonucleotides are useful for inhibition of NEK expression in normal and tumor cells in order to profile the potential effects of small molecule NEK inhibitors. They can also be used to inhibit NEK expression in human tumor cell xenografts in nude mice to determine the antitumor effects of NEK inhibitors. Another use is as a "drug" to inhibit NEK expression in various human tumors that are characterized by overexpression, or aberrant expression, of a *nek* gene.

Material and Methods:

Antisense oligonucleotides to be tested for their inhibition of *nek* expression can be generated by methods well-known in the art. Typically, the oligonucleotides are synthesized as phosphorothioates, and are tested in the following manner.

Human H1299 cells are seeded at ~40-50% confluency in a 6 well plate (Falcon). The following day lipofectin (Gibco) and oligo(s) are mixed with OptiMEM (Gibco) such that the final concentration of lipofectin is 100 μg/mL and the final concentration of each oligo is 1 μM in a volume of 200 μL. The lipofectin/oligo/OptiMEM mixture is incubated at room temperature (20-25 °C) for 15 minutes, 800 μL of OptiMEM is added to the lipofectin/oligo/OptiMEM mixture, and is mixed gently. The growth medium is removed from the H1299 cells, which are at ~80% confluency. Cells are washed once with OptiMEM. The OptiMEM is aspirated, the lipofectin/oligo/OptiMEM mixture is added, and cultures are incubated at 37 °C for 4 hours. The lipofectin/oligo/OptiMEM mixture is removed and is replaced with normal growth medium containing the antisense oligo(s) at a concentration of 200 nM. The plates are returned to the 37 °C incubator for 16-20 hours after which time the growth medium is removed from the cells, and they are washed once with OptiMEM. The OptiMEM is again aspirated, and the lipofectin/oligo/OptiMEM mixture is added (prepared as described above), and the plates again are incubated at 37 °C for 4 hours. The lipofectin/oligo/OptiMEM mixture is removed again, and is replaced with normal growth medium containing the antisense oligo(s) at a concentration of 200 nM. The plates are returned to the 37 °C incubator for 16-20 hours. The cells are harvested (~40 hours after initial treatment), and are analyzed for *nek* mRNA by Northern blot or NEK protein expression by immunoblot.

One skilled in the art would readily appreciate that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. The molecular complexes and the methods, procedures, treatments, molecules, specific compounds described herein are presently representative of preferred embodiments are exemplary and are not intended as limitations on the scope of the invention. Changes therein and other uses will occur to those skilled in the art which are encompassed within the spirit of the invention are defined by the scope of the claims.

It will be readily apparent to one skilled in the art that varying substitutions and modifications may be made to the invention disclosed herein without departing from the scope and spirit of the invention.

All patents and publications mentioned in the specification are indicative of the levels of those skilled in the art to which the invention pertains.

The invention illustratively described herein suitably may be practiced in the absence of any element or elements, limitation or limitations which is not specifically disclosed herein. Thus, for example, in each instance herein any of the terms "comprising", "consisting essentially of" and "consisting of" may be replaced with either of the other two terms. The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed.

In particular, although some formulations described herein have been identified by the excipients added to the formulations, the invention is meant to also cover the final formulation formed by the combination of these excipients. Specifically, the invention includes formulations in which one to all of the added excipients undergo a reaction during formulation and are no longer present in the final formulation, or are present in modified forms.

In addition, where features or aspects of the invention are described in terms of Markush groups, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group. For example, if X is described as selected from the group consisting of bromine, chlorine, and iodine, claims for X being bromine and claims for X being bromine and chlorine are fully described.

Other embodiments are within the following claims.

What is claimed is:

CLAIMS

1. An isolated, enriched or purified nucleic acid molecule encoding a BUB1 kinase polypeptide, a NEK5 kinase polypeptide or a NEK6 kinase polypeptide, wherein said nucleic acid molecule comprises a nucleotide sequence that:

(a) encodes a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6;

(b) is the complement of the nucleotide sequence of (a);

(c) hybridizes under stringent conditions to the nucleotide molecule of (a) and encodes a naturally occurring kinase polypeptide;

(d) encodes a kinase polypeptide having the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6, except that it lacks one or more, but not all, of the following segments of amino acid residues: 1-785, 786-1028, or 1029-1085 of SEQ ID NO:2; 1-31, 32-283, or 284-302 of SEQ ID NO:4 or 1-22 or 23-302 of SEQ ID NO:6;

(e) is the complement of the nucleotide sequence of (d);

(f) encodes a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6 from amino acid residues 1-785, 786-1028, or 1029-1085 of SEQ ID NO:2; 1-31, 32-283, or 284-302 of SEQ ID NO:4 or 1-22 or 23-302 of SEQ ID NO:6;

(g) is the complement of the nucleotide sequence of (f);

(h) encodes a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6, except that it lacks one or more, but not all, of an N-terminal domain, a C-terminal catalytic domain, a catalytic domain, a C-terminal domain, and a C-terminal tail; or

(i) is the complement of the nucleotide sequence of (h).

2. The nucleic acid molecule of claim 1, further comprising a vector or promoter effective to initiate transcription in a host cell.

3. The nucleic acid molecule of claim 1, wherein said nucleic acid molecule is isolated, enriched, or purified from a mammal.

4. The nucleic acid molecule of claim 3, wherein said mammal is a human.

5. The nucleic acid probe of claim 1 used for the detection of nucleic acid encoding a kinase polypeptide in a sample, wherein said kinase polypeptide is selected from the group consisting of a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2, a NEK5 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:4 and a NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6.

6. A recombinant cell comprising the nucleic acid molecule of claim 1 encoding a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2, a NEK5 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:4 or a NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6.

7. An isolated, enriched, or purified BUB1 kinase polypeptide NEK5 kinase polypeptide or NEK6 kinase polypeptide, wherein said polypeptide comprises an amino acid sequence having

(a) the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6, respectively;

(b) the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6, respectively, except that it lacks one or more, but not all, of the following segments of amino acid residues: 1-785, 786-1028, or 1029-1085 of SEQ ID NO:2; 1-31, 32-283, or 284-302 of SEQ ID NO:4; or 1-22 or 23-302 of SEQ ID NO:6;

(c) the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6, respectively, from amino acid residues 1-785, 786-1028, or 1029-1085 of SEQ ID NO:2; 1-31, 32-283, or 284-302 of SEQ ID NO:4; or 1-22 or 23-302 of SEQ ID NO:6; or

(d) the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6, respectively, except that it lacks one or more, but not all, of the domains selected from the group consisting of an N-terminal domain, a C-terminal catalytic domain, a catalytic domain, a C-terminal domain, and a C-terminal tail.

8. The kinase polypeptide of claim 7, wherein said polypeptide is isolated, purified, or enriched from a mammal.

9. The kinase polypeptide of claim 8, wherein said mammal is a human.

10. The kinase polypeptide of claim 7, wherein said polypeptide is a BUB1 kinase polypeptide as set forth in SEQ ID NO:2.

11. The kinase polypeptide of claim 7, wherein said polypeptide is a NEK5 kinase polypeptide as set forth in SEQ ID NO:4 or a NEK6 kinase polypeptide as set forth in SEQ ID NO:6.

12. An antibody or antibody fragment having specific binding affinity to a kinase polypeptide or to a domain of said polypeptide, wherein said polypeptide is a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2, a NEK 4a kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:8, a NEK 4b kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:10, a NEK5 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:4 or a NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6.

13. A hybridoma which produces an antibody having specific binding affinity to a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2, a NEK 4a kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:8, a NEK 4b kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:10, a NEK5 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:4 or a NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6.

14. A method for identifying a substance that modulates the activity of a BUB1 kinase polypeptide, a NEK4a kinase polypeptide, a NEK4b kinase polypeptide, a NEK5 kinase polypeptide activity, or a NEK6 kinase polypeptide comprising the steps of:

(a) contacting the BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2, NEK4a kinase polypeptide having the amino acid

sequence set forth in SEQ ID NO:8, NEK4b kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:10, NEK5 kinase polypeptide activity having the amino acid sequence set forth in SEQ ID NO:4, or NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6 with a test substance;

(b) measuring the activity of said polypeptide; and

(c) determining whether said substance modulates the activity of said polypeptide.

15. A method for identifying a substance that modulates the activity of a BUB1 kinase polypeptide, a NEK4a kinase polypeptide, a NEK4b kinase polypeptide, a NEK5 kinase polypeptide, or a NEK6 kinase polypeptide in a cell comprising the steps of:

(a) expressing a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2, a NEK4a kinase polypeptide having the amino acid sequence forth in SEQ ID NO:8, a NEK4b kinase polypeptide having the amino acid sequence forth in SEQ ID NO:10, a NEK5 kinase polypeptide activity having the amino acid sequence set forth in SEQ ID NO:4, or a NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6 in a cell;

(b) adding a test substance to said cell; and

(c) monitoring a change in cell phenotype or the interaction between said polypeptide and a natural binding partner.

16. A method for treating a disease or a disorder comprising administering to a patient in need of such treatment a substance that modulates the activity of a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2.

17. The method of claim 18, wherein said disease is cancer.

18. The method of claim 19, wherein said cancer is selected from the group consisting of leukemia, cervical cancer, lymphoma, colon cancer, lung cancer, melanoma, ovarian cancer, CNS cancer, prostate cancer, kidney cancer, and breast cancer.

19. The method of claim 18, wherein said substance modulates kinase activity *in vitro*.

20. A method for detection of nucleic acid encoding a BUB1 kinase polypeptide as set forth in SEQ ID NO:2 in a sample as a diagnostic tool for a disease or a disorder, wherein said method comprises the steps of:

(a) contacting said sample with a nucleic acid probe which hybridizes under hybridization assay conditions to a nucleic acid target region of *bub1* having the amino acid sequence set forth in SEQ ID NO:1, said probe comprising the nucleic acid sequence encoding said polypeptide, fragments thereof, or the complements of said sequences and fragments; and

(b) detecting the presence or amount of the probe:target region hybrid as an indication of said disease.

21. The method of claim 20, wherein said disease is cancer.

22. The method of claim 21, wherein said cancer is selected from the group consisting of leukemia, cervical cancer, lymphoma, colon cancer, lung cancer, melanoma, ovarian cancer, CNS cancer, prostate cancer, kidney cancer, and breast cancer.

23. A method for treating a disease or a disorder comprising administering to a patient in need of such treatment a substance that modulates the activity of a NEK4a kinase polypeptide having the amino acid sequence forth in SEQ ID NO:8, a NEK4b kinase polypeptide having the amino acid sequence forth in SEQ ID NO:10, a NEK5 kinase polypeptide activity having the amino acid sequence set forth in SEQ ID NO:4, or a NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6.

24. The method of claim 23, wherein said disease or disorder is selected from the group consisting of cancer, neurodegenerative, and immune disorders.

25. The method of claim 24, wherein said disease or disorder is cancer.

26. The method of claim 23, wherein said substance modulates the *in vitro* activity of the NEK4a kinase polypeptide, the NEK4b kinase polypeptide, the NEK5 kinase polypeptide, or the NEK6 kinase polypeptide.

27. The method of claim 26, wherein said wherein said modulation is inhibition.

28. The method of claim 27, wherein said substance is an antisense oligonucleotide.

29. A method for detection of nucleic acid encoding a kinase polypeptide in a sample as a diagnostic tool for a disease or a disorder, comprising the steps of:

(a) contacting said sample with a nucleic acid probe which hybridizes under hybridization assay conditions to a nucleic acid target region of a kinase polypeptide, said probe comprising the nucleic acid sequence encoding a BUB 1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2, a NEK 5 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:4, a NEK 6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6, a Nek 4a kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:8, or a NEK 4b kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:10, fragments thereof, and the complements of said sequences and fragments; and

(b) detecting the presence or amount of the probe:target region hybrid as an indication of said disease.

30. The method of claim 29, wherein said disease is selected from the group consisting of cancer, neurodegenerative, and immune disorders.

31. The method of claim 30, wherein said disease is cancer.

32. An antisense oligonucleotide that inhibits the expression of a BUB1 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:2, a NEK4a kinase polypeptide having the amino acid sequence forth in SEQ ID NO:8, a NEK4b kinase polypeptide having the amino acid sequence forth in SEQ ID NO:10, a NEK5 kinase polypeptide activity having the amino acid sequence set forth in SEQ ID NO:4, or a NEK6 kinase polypeptide having the amino acid sequence set forth in SEQ ID NO:6

33. The antisense oligonucleotide of claim 34, wherein said oligonucleotide is the complement of a sequence encoding a fragment of the protein encoded by the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10.

BUB1 human nucleotide sequence

GTTTGCGGTTCAGGTTTGGCCGCTGCCGGCCAGCGTCCTCTGGCCATGGACACCCCGGA
AAATGTCCTTCAGATGCTTGAAGCCCACATGCAGAGCTACAAGGGCAATGACCCTCTTG
GTGAATGGGAAAGATACATACAGTGGGTAGAAGAGAATTTTCCTGAGAATAAAGAATAC
TTGATAACTTTACTAGAACATTTAATGAAGGAATTTTTAGATAAGAAGAAATACCACAA
TGACCCAAGATTCATCAGTTATTGTTTAAAATTTGCTGAGTACAACAGTGACCTCCATC
AATTTTTTGAGTTTCTGTACAACCATGGGATTGGAACCCTGTCATCCCCTCTGTACATT
GCCTGGGCGGGGCATCTGGAAGCCCAAGGAGAGCTGCAGCATGCCAGTGCTGTCCTTCA
AAGAGGAATTCAAAACCAGGCTGAACCCAGAGAGTTCCTGCAACAACAATACAGGTTAT
TTCAGACACGCCTCACTGAAACCCATTTGCCAGCTCAAGCTAGAACCTCAGAACCTCTG
CATAATGTTCAGGTTTTAAATCAAATGATAACATCAAAATCAAATCCAGGAAATAACAT
GGCCTGCATTTCTAAGAATCAGGGTTCAGAGCTTTCTGGAGTGATATCTTCAGCTTGTG
ATAAAGAGTCAAATATGGAACGAAGAGTGATCACGATTTCTAAATCAGAATATTCTGTG
CACTCATCTTTGGCATCCAAAGTTGATGTTGAGCAGGTTGTTATGTATTGCAAGGAGAA
GCTTATTCGTGGGGAATCAGAATTTTCCTTTGAAGAATTGAGAGCCCAGAAATACAATC
AACGGAGAAAGCATGAGCAATGGGTAAATGAAGACAGACATTATATGAAAAGGAAAGAA
GCAAATGCTTTTGAAGAACAGCTATTAAAACAGAAAATGGATGAACTTCATAAGAAGTT
GCATCAGGTGGTGGAGACATCCCATGAGGATCTGCCCGCTTCCCAGGAAAGGTCCGAGG
TTAATCCAGCACGTATGGGGCCAAGTGTAGGCTCCCAGCAGGAACTGAGAGCGCCATGT
CTTCCAGTAACCTATCAGCAGACACCAGTGAACATGGAAAAGAACCCAAGAGAGGCACC
TCCTGTTGTTCCTCCTTTGGCAAATGCTATTTCTGCAGCTTTGGTGTCCCCAGCCACCA
GCCAGAGCATTGCTCCTCCTGTTCCTTTGAAAGCCCAGACAGTAACAGACTCCATGTTT
GCAGTGGCCAGCAAAGATGCTGGATGTGTGAATAAGAGTACTCATGAATTCAAGCCACA
GAGTGGAGCAGAGATCAAAGAAGGGTGTGAAACACATAAGGTTGCCAACACAAGTTCTT
TTCACACAACTCCAAACACATCACTGGGAATGGTTCAGGCAACGCCATCCAAAGTGCAG

FIG. 1A

CCATCACCCACCGTGCACACAAAAGAAGCATTAGGTTTCATCATGAATATGTTTCAGGC
TCCTACACTTCCTGATATTTCTGATGACAAAGATGAATGGCAATCTCTAGATCAAAATG
AAGATGCATTTGAAGCCCAGTTTCAAAAAAATGTAAGGTCATCTGGGGCTTGGGGAGTC
AATAAGATCATCTCTTCTTTGTCATCTGCTTTTCATGTGTTTGAAGATGGAAACAAAGA
AAATTATGGATTACCACAGCCTAAAAATAAACCCACAGGAGCCAGGACCTTTGGAGAAC
GCTCTGTCAGCAGACTTCCTTCAAAACCAAAGGAGGAAGTGCCTCATGCTGAAGAGTTT
TTGGATGACTCAACTGTATGGGGTATTCGCTGCAACAAAACCCTGGCACCCAGTCCTAA
GAGCCCAGGAGACTTCACATCTGCTGCACAACTTGCGTCTACACCATTCCACAAGCTTC
CAGTGGAGTCAGTGCACATTTTAGAAGATAAAGAAAATGTGGTAGCAAAACAGTGTACC
CAGGCGACTTTGGATTCTTGTGAGGAAAACATGGTGGTGCCTTCAAGGGATGGAAAATT
CAGTCCAATTCAAGAGAAAAGCCCAAAACAGGCCTTGTCGTCTCACATGTATTCAGCAT
CCTTACTTCGTCTGAGCCAGCCTGCTGCAGGTGGGGTACTTACCTGTGAGGCAGAGTTG
GGCGTTGAGGCTTGCAGACTCACAGACACTGACGCTGCCATTGCAGAAGATCCACCAGA
TGCTATTGCTGGGCTCCAAGCAGAATGGATGCAGATGAGTTCACTTGGGACTGTTGATG
CTCCAAACTTCATTGTTGGGAACCCATGGGATGATAAGCTGATTTTCAAACTTTTATCT
GGGCTTTCTAAACCAGTGAGTTCCTATCCAAATACTTTTGAATGGCAATGTAAACTTCC
AGCCATCAAGCCCAAGACTGAATTTCAATTGGGTTCTAAGCTGGTCTATGTCCATCACC
TTCTTGGAGAAGGAGCCTTTGCCCAGGTGTACGAAGCTACCCAGGGAGATCTGAATGAT
GCTAAAAATAAACAGAAATTTGTTTTAAAGGTCCAAAAGCCTGCCAACCCCTGGGAATT
CTACATTGGGACCCAGTTGATGGAAAGACTAAAGCCATCTATGCAGCACATGTTTATGA
AGTTCTATTCTGCCCACTTATTCCAGAATGGCAGTGTATTAGTAGGAGAGCTCTACAGC
TATGGAACATTATTAAATGCCATTAACCTCTATAAAAATACCCCTGAAAAAGTGATGCC
TCAAGGTCTTGTCATCTCTTTTGCTATGAGAATGCTTTACATGATTGAGCAAGTGCATG
ACTGTGAAATCATTCATGGAGACATTAAACCAGACAATTTCATACTTGGAAACGGATTT
TTGGAACAGGATGATGAAGATGATTTATCTGCTGGCTTGGCACTGATTGACCTGGGTCA

FIG. 1B

GAGTATAGATATGAAACTTTTTCCAAAAGGAACTATATTCACAGCAAAGTGTGAAACAT
CTGGTTTTCAGTGTGTTGAGATGCTCAGCAACAAACCATGGAACTACCAGATCGATTAC
TTTGGGGTTGCTGCAACAGTATATTGCATGCTCTTTGGCACTTACATGAAAGTGAAAAA
TGAAGGAGGAGAGTGTAAGCCTGAAGGTCTTTTTAGAAGGCTTCCTCATTTGGATATGT
GGAATGAATTTTTTCATGTTATGTTGAATATTCCAGATTGTCATCATCTTCCATCTTTG
GATTTGTTAAGGCAAAAGCTGAAGAAAGTATTTCAACAACACTATACTAACAAGATTAG
GGCCCTACGTAATAGGCTAATTGTACTGCTCTTAGAATGTAAGCGTTCACGAAAATAAA
ATTTGGATATAGACAGTCCTTAAAAATCACACTGTAAATATGAATCTGCTCACTTTAAA
CCTGTTTTTTTTTCATTTATTGTTTATGTAAATGTTTGTCAAAAA

FIG. 1C

BUB1 aa sequence

MDTPENVLQMLEAHMQSYKGNDPLGEWERYIQWVEENFPENKEYLITLLE
HLMKEFLDKKKYHNDPRFISYCLKFAEYNSDLHQFFEFLYNHGIGTLSSP
LYIAWAGHLEAQGELQHASAVLQRGIQNQAEPREFLQQQYRLFQTRLTET
HLPAQARTSEPLHNVQVLNQMITSKSNPGNNMACISKNQGSELSGVISSA
CDKESNMERRVITISKSEYSVHSSLASKVDVEQVVMYCKEKLIRGESEFS
FEELRAQKYNQRRKHEQWVNEDRHYMKRKEANAFEEQLLKQKMDELHKKL
HQVVETSHEDLPASQERSEVNPARMGPSVGSQQELRAPCLPVTYQQTPVN
MEKNPREAPPVVPPLANAISAALVSPATSQSIAPPVPLKAQTVTDSMFAV
ASKDAGCVNKSTHEFKPQSGAEIKEGCETHKVANTSSFHTTPNTSLGMVQ
ATPSKVQPSPTVHTKEALGFIMNMFQAPTLPDISDDKDEWQSLDQNEDAF
EAQFQKNVRSSGAWGVNKIISSLSSAFHVFEDGNKENYGLPQPKNKPTGA
RTFGERSVSRLPSKPKEEVPHAEEFLDDSTVWGIRCNKTLAPSPKSPGDF
TSAAQLASTPFHKLPVESVHILEDKENVVAKQCTQATLDSCEENMVVPSR
DGKFSPIQEKSPKQALSSHMYSASLLRLSQPAAGGVLTCEAELGVEACRL
TDTDAAIAEDPPDAIAGLQAEWMQMSSLGTVDAPNFIVGNPWDDKLIFKL
LSGLSKPVSSYPNTFEWQCKLPAIKPKTEFQLGSKLVYVHHLLGEGAFAQ
VYEATQGDLNDAKNKQKFVLKVQKPANPWEFYIGTQLMERLKPSMQHMFM
KFYSAHLFQNGSVLVGELYSYGTLLNAINLYKNTPEKVMPQGLVISFAMR
MLYMIEQVHDCEIIHGDIKPDNFILGNGFLEQDDEDDLSAGLALIDLGQS
IDMKLFPKGTIFTAKCETSGFQCVEMLSNKPWNYQIDYFGVAATVYCMLF
GTYMKVKNEGGECKPEGLFRRLPHLDMWNEFFHVMLNIPDCHHLPSLDLL
RQKLKKVFQQHYTNKIRALRNRLIVLLLECKRSRK

FIG. 2

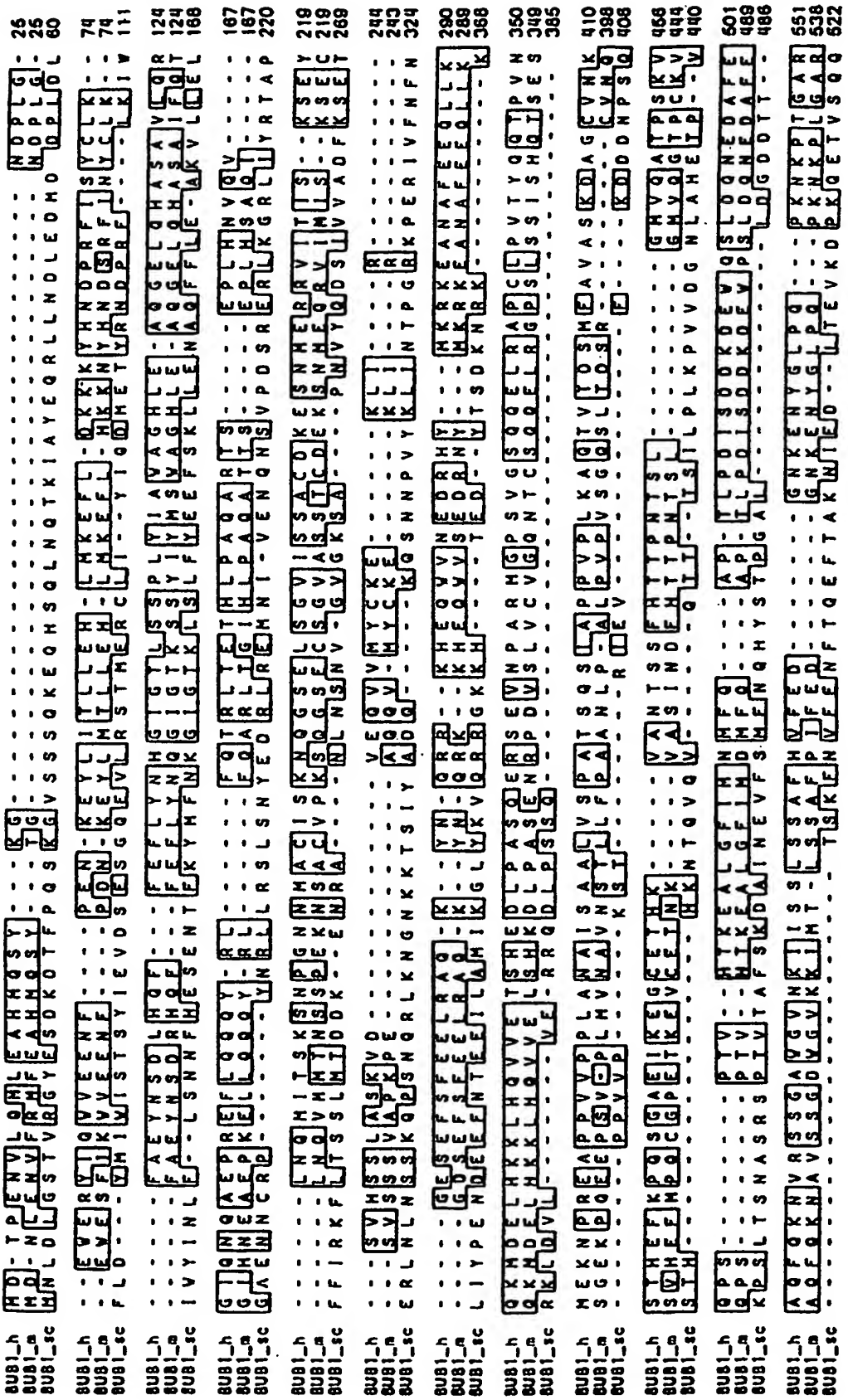

FIG. 3A

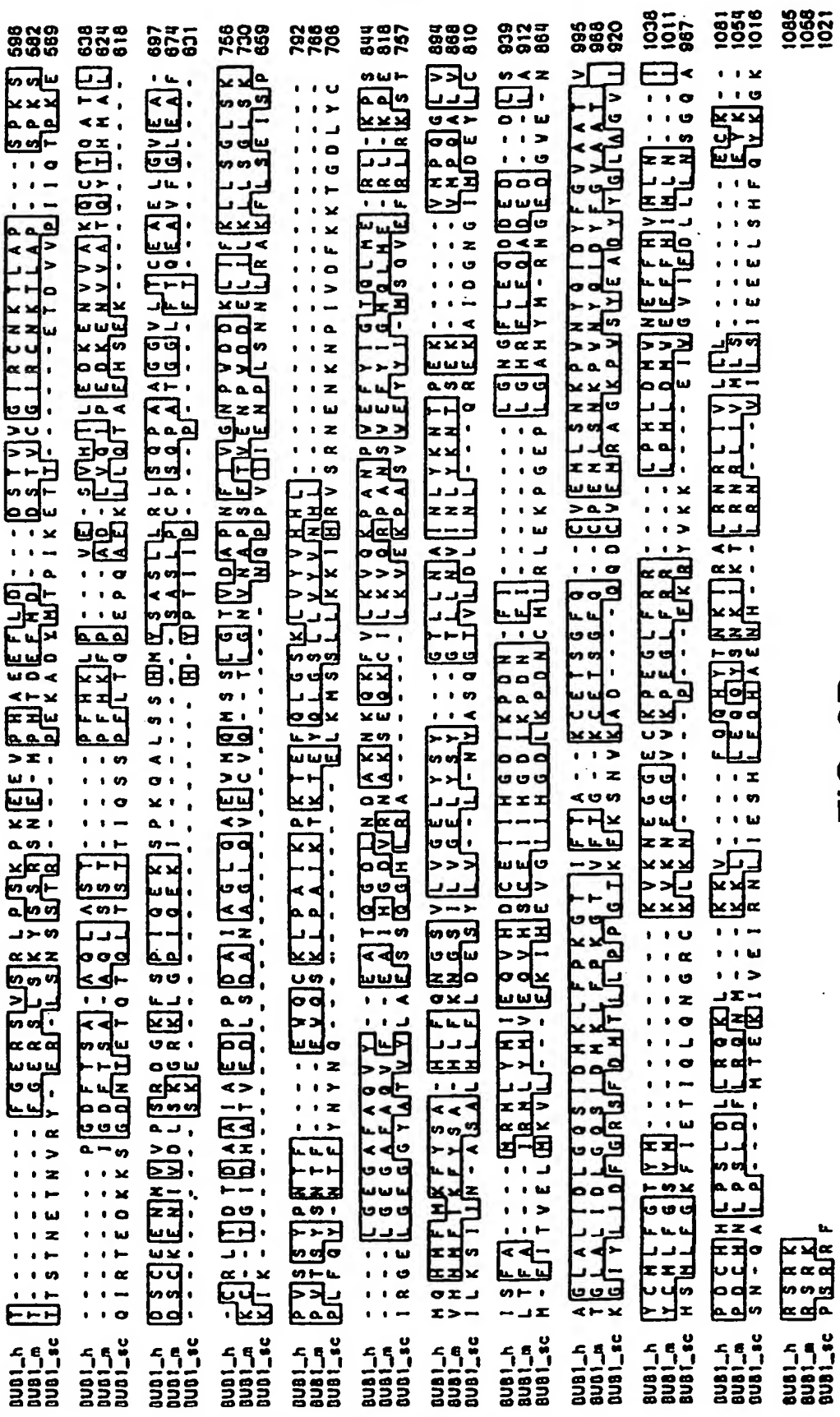

FIG. 3B

Human NEK4a

TGCGCCGCAAACTCGTGTGGGACGCACCGCTCCAGCCGCCCGCGGGCCAGCGCACCGGTC
CCCCAGCGGCAGCCGAGCCCGCCCGCGCGCCGTTCGTGCCCTCGTGAGGCTGGCATGCAG
GATGGCAGGACAGCCCGGCCACATGCCCCATGGAGGGAGTTCCAACAACCTCTGCCACAC
CCTGGGGCCTGTGCATCCTCCTGACCCACAGAGGCATCCCAACACGCTGTCTTTTCGCTG
CTCGCTGGCGGACTTCCAGATCGAAAAGAAGATAGGCCGAGGACAGTTCAGCGAGGTGTA
CAAGGCCACCTGCCTGCTGGACAGGAAGACAGTGGCTCTGAAGAAGGTGCAGATCTTTGA
GATGATGGACGCCAAGGCGAGGCAGGACTGTGTCAAGGAGATCGGCCTCTTGAAGCAACT
GAACCACCCAAATATCATCAAGTATTTGGACTCGTTTATCGAAGACAACGAGCTGAACAT
TGTGCTGGAGTTGGCTGACGCAGGGGACCTCTCGCAGATGATCAAGTACTTTAAGAAGCA
GAAGCGGCTCATCCCGGAGAGGACAGTATGGAAGTACTTTGTGCAGCTGTGCAGCGCCGT
GGAGCACATGCATTCACGCCGGGTGATGCACCGAGACATCAAGCCTGCCAACGTGTTCAT
CACAGCCACGGGCGTCGTGAAGCTCGGTGACCTTGGTCTGGGCCGCTTCTTCAGCTCTGA
GACCACCGCAGCCCACTCCCTAGTGGGGACGCCCTACTACATGTCACCGGAGAGGATCCA
TGAGAACGGCTACAACTTCAAGTCCGACATCTGGTCCTTGGGCTGTCTGCTGTACGAGAT
GGCAGCCCTCCAGAGCCCCTTCTATGGAGATAAGATGAATCTCTTCTCCCTGTGCCAGAA
GATCGAGCAGTGTGACTACCCCCCACTCCCCGGGGAGCACTACTCCGAGAAGTTACGAGA
ACTGGTCAGCATGTGCATCTGCCCTGACCCCCACCAGAGACCTGGTAGCCTAGAACAGCT
AAGACCACAGGGTTCAGCAGGTTCCCCAAAAGGCTGCCCAGCCTTACAGCAGATGCTGAA
GGCAGAGCAGCTGAGGGAGGGGCGCTGGCCACATGTCACTGATGGTCAGATTCCAAAGTC
CTTTCTTTATACTGTTGTGGACAATCTCAGCTGGGTCAATAAGGGCAGGTGGTTCAGCGA
GCCACGGCAGCCCCCTGTATCTGGATTGTAATGTGAATCTTTAGGGTAATTCCTCCAGTG
ACCTGTCAAGGCTTATGCTAACAGGAGACTTGCAGGAGACCGTGTGATTTGTGTAGTGAG
CCTTTGAAAATGGTTAGTACCGGGTTCAGTTTAGTTCTTGGTATCTTTTCAATCAAGCTG
TGTGCTTAATTTACTCTGTTGTAAAGGGATAAAGTGGAAATCATTTTTAAAAAAAAAAAA
AAAAAAAAA

Human NEK4b

TCCAGATCGAAAAGAAGATAGGCCGAGGACAGTTCAGCGAGGTGTACAAGGCCACCTGCC
TGCTGGACAGGAAGACAGTGGCTCTGAAGAAGGTGCAGATCTTTGAGATGATGGACGCCA
AGGCGAGGCAGGACTGTGTCAAGGAGATCGGCCTCTTGAAGCAACTGAACCACCCAAATA
TCATCAAGTATTTGGACTCGTTTATCGAAGACAACGAGCTGAACATTGTGCTGGAGTTGG
CTGACGCAGGGGACCTCTCGCAGATGATCAAGTACTTTAAGAAGCAGAAGCGGCTCATCC
CGGAGAGGACAGTATGGAAGTACTTTGTGCAGCTGTGCAGCGCCGTGGAGCACATGCATT
CACGCCGGGTGATGCACCGAGACATCAAGCCTGCCAACGTGTTCATCACAGCCACGGGCG
TCGTGAAGCTCGGTGACCTTGGTCTGGGCCGCTTCTTCAGCTCTGAGACCACCGCAGCCC
ACTCCCTAGTGGGGACGCCCTACTACATGTCACCGGAGAGGATCCATGAGAACGGCTACA
ACTTCAAGTCCGACATCTGGTCCTTGGGCTGTCTGCTGTACGAGATGGCAGCCCTCCAGA
GCCCCTTCTATGGAGATAAGATGAATCTCTTCTCCCTGTGCCAGAAGATCGAGCAGTGTG
ACTACCCCCCACTCCCCGGGGAGCACTACTCCGAGAAGTTACGAGAACTGGTCAGCATGT
GCATCTGCCCTGACCCCCACCAGAGACCTGACATCGGATACGTGCACCAGGTGGCCAAGC
AGATGCACATCTGGATGTCCAGCACCTGAGCGTGGATGCACCGTGCCTTATCAAAGCCAG
CACCACTTTGCCTTACTTGAGTCGTCTTCTCTTCGAGTGGCCACCTGGTAGCCTAGAACA
GCTAAGACCACAGGGTTCAGCAGGTTCCCCAAAAGGCTGCCCAGCCTTACAGCAGATGCT
GAAGGCAGAGCAGCTGAGGGAGGGGCGCTGGCCACATGTCACTGATGGTCAGATTCCAAA

FIG. 4A

GTCCTTTCTTTATACTGTTGTGGACAATCTCAGCTGGGTCAATAAGGGCAGGTGGTTCAG
CGAGCCACGGCAGCCCCCTGTATCTGGATTGTAATGTGAATCTTTAGGGTAATTCCTCCA
GTGACCTGTCAAGGCTTATGCTAACAGGAGACTTGCAGGAGACCGTGTGATTTGTGTAGT
GAGCCTTTGAAAATGGTTAGTACCGGGTTCAGTTTAGTTCTTAGTATCTTTTCAATCAAG
CTGTGTGCTTAATTTACTCT
GTTGTAAAGGGATAAAGTGGAAATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAA

Human NEK5
CTGAGTTCTAAAGTTCCTGTTGCTTCAGACAATGGATGAGCAATCACAAGGAATGCAAGG
GCCACCTGTTCCTCAGTTCCAACCACAGAAGGCCTTACGACCGGATATGGGCTATAATAC
ATTAGCCAACTTTCGAATAGAAAAGAAAATTGGTCGCGGACAATTTAGTGAAGTTTATAG
AGCAGCCTGTCTCTTGGATGGAGTACCAGTAGCTTTAAAAAAAGTGCAGATATTTGATTT
AATGGATGCCAAAGCACGTGCTGATTGCATCAAAGAAATAGATCTTCTTAAGCAACTCAA
CCATCCAAATGTAATAAAATATTATGCATCATTCATTGAAGATAATGAACTAAACATAGT
TTTGGAACTAGCAGATGCTGGCGACCTATCCAGAATGATCAAGCATTTTAAGAAGCAAAA
GAGGCTAATTCCTGAAAGAACTGTTTGGAAGTATTTTGTTCAGCTTTGCAGTGCATTGGA
ACACATGCATTCTCGAAGAGTCATGCATAGAGATATAAAACCAGCTAATGTGTTCATTAC
AGCCACTGGGGTGGTAAAACTTGGAGATCTTGGGCTTGGCCGGTTTTTCAGCTCAAAAAC
CACAGCTGCACATTCTTTAGTTGGTACGCCTTATTACATGTCTCCAGAGAGAATACATGA
AAATGGATACAACTTCAAATCTGACATCTGGTCTCTTGGCTGTCTACTATATGAGATGGC
TGCATTACAAAGTCCTTTCTATGGTGACAAAATGAATTTATACTCACTGTGTAAGAAGAT
AGAACAGTGTGACTACCCACCTCTTCCTTCAGATCACTATTCAGAAGAACTCCGACAGTT
AGTTAATATGTGCATCAACCCAGATCCAGAGAAGCGACCAGACGTCACCTATGTTTATGA
CGTAGCAAAGAGGATGCATGCATGCACTGCAAGCAGCTAAACATGCAAGATCATGAAGAG
TGTAACCAAAGTAATTGAAAGTATTTTGTGCAAGTCATACCTCCCCATTTATGTCTGGTG
TTAAGATTAATATTTCAGAGCTAGTGTGCTTTGAATCCTTAACCAGTTTTCATATAAGCT
TCATTTTGTACCAGTCACCTAAATCACCTCCTTGCAACCCCCAAATGACTTTGGAATAAC
TGAATTGCATGTTAGGAGAGAAAATGAAACATGATGGTTTTGAATGGCTAAAGGTTTATA
GAATTTCTTACAGTTTTCTGCTGATAAATTGTGTTTAGATAGACTGTCAGTGCCAAATAT
TGAAGGTGCAGCTTGGCACACATCAGAATAGACTCATACCTGAGAAAAAGTATCTGAACA
TGTGACTTGTTTCTTTTTTAGTAATTTATGGACATTGAGATGAACACAATTGTGAACTTT
TGTGAAGATTTTATTTTTAAACGTTTGAAGTACTAGTTTTAGTTCTTAGCAGAGTAGTTT
TCAAATATGATTCTTATGATAAATGTAGACACAAACTATTTGAGAAACATTTAGAACTCT
TAGCTTATACATTCAAAATGTAACTATTAAATGTGAAGATTTGGGGACAAAATGTGAGTC
AGACACTGAAGAGTTTTTTGTTTTGTTTTAATATTTTTGATATTCTCTTTGCATTGAAAT
GGTATAAATGAATCCATTTAAAAAGTGGTTAAGGATTTGTTTAGCTGGTGTGATAATAAT
TTTTAAAGTTGCACATTGCCCAAGGCTTTTTTTGTGTGTTTTTATTGTTGTTTGTACATT
TGAAAAATATTCTTTGAATAACCTTGCAGTACTATATTTCAATTTCTTTATAAATTTAAG
TGCATTTTAACTCATAATTGTACACTATAATATAAGCCTAAGTTTTTATTCATAAGTTTT
ATTGAAGTTCTGATCGGTCCCCTTCAGAAATTTTTTTATATTATTCTTCAAGTTACTTTC
TTATTTATATTGTATGTGCATTTTATCCATTAATGTTTCATACTTTCTGAGAGTATAATA
CCCTTTTAAAAGATATTTGGTATACCAATACTTTTCCTGGATTGAAAACTTTTTTTAAAC
TTTTTAAAATTTGGGCCACTCTGTATGCATATGTTTGGTCTTGTTAAAGAGGAAGAAAGG
ATGTGTGTTATACTGTACCTGTGAATGTTGATACAGTTACAATTTATTTGACAAGGTTGT

FIG. 4B

AATTCTAGAATATGCTTAATAAAATGAAAACTGGCCATGACTACAGCCAGAACTGTTATG
AGATTAACATTTCTATTGAGAAGCTTTTGAGTAAAGTACTGTATTTGTTCATGAAGATGA
CTGAGATGGTAACACTTCGTGTAGCTTAAGGAAATGGGCAGAATTTCGTAAATGCTGTTG
TGCAGATGTGTTTTCCCTGAATGCTTTCGTATTAGTGGCGACCAGTTTCTCACAGAATTG
TGAAGCCTGAAGGCCAAGAGGAAGTCACTGTTAAAGGACTCTGTGCCATCTTACAACCTT
GGATGAATTATCCTGCCAACGTGAAAACCTCATGTTCAAAGAACACTTCCCTTTAGCCGA
TGTAACTGCTGGTTTTGTTTTTCATATGTGTTTTTCTTACACTCATTTGAATGCTTTCAA
GCATTTGTAAACTTAAAAAATGTATAAAGGGCAAAAAGTCTGAACCCTTGTTTTCTGAAA
TCTAATCAGTTATGTATGGTTTCTGAAGGGTAATTTTATTTTGGAATAGGTAAAGGAAAC
CTGTTTTGTTTGTTTTTCCTGAGGGCTAGATGCATTTTTTTTCTCACACTCTTAATGACT
TTTAACATTTATACTGAGCATCCATAGATATATTCCTAGAAGTATGAGAAGAATTATTCT
TATTGACCATTAATGTCATGTTCATTTTAATGTAATATAATTGAGATGAAATGTTCTCTG
GTTGGAACAGATACTCTCTTTTTTTTCTTGCAATCTTTAAGAATACATAGATCTAAAATT
CATTAGCTTGACCCCTCAAAGTAACTTTTAAGTAAAGATTAAAGCTTTTCTTCTCAGTGA
ATATATCTGCTAGAAGGAAATAGCTGGGAAGAATTTAATGATCAGGGAAATTCATTATTT
CTATATGTGGAAACTTTTTGCTTCGAATATTGTATCTTTTTAAATCTAAATGTTCATATT
TTTCCTGAAGAAACCACTGTGTAAAAATCAAATTTTAATTTTGAATGGAATAATTTCAAA
GAACTATGAAGATGATTTGAAGCTCTAATTTATATAGTCACCTATAAAATGTTCTTTATA
TGTGTTCATAAGTAAATTTTATATTGATTAAGTTAAACTTTTGAATTGATTTGAGGAGCA
GTAAAATGAAAGCTATATCTATTCTAAACCTTATTTAGACATTGGTACCAGTTACCCAGG
TGAAAATATGGAGTAACTTTGTTTTGTATGGTAAGGTTTAGGAATGGTGGATGAAGGGTA
TCTCTATATAAATAAAGTGCTCAACAATGTGCAATGATTGTAAATTTAGTAAGATATTAC
AGCCATTTCATGAATGCTTTACCATTCAACATAGTATCTATTACAAAACACCTTTCTTGT
ATCCATATACTTCAGGTGTTGCTGTTAACATTTACTATGATATTTATTTTAACCAAAATG
TTACTCACATTAAATGTTTATTCTTTAAAATGAATGTATTATGTTTTTAACCCACAAATG
CATACTTACCCTGTGCCTCATATTTCAATAGTACTGTAATATGGACATCTTTTGTGAAAT
ACTTTTATTTTGTTATGCTTTAAATATACATACAAAAAGATTTCTGTTATTAGCTTTGAA
AATTGTATAATATCCTAATATAAACAAAAATATAAAAATAAAAATGAATACAGTAAAAAA
AAAAAAAAAAAAA

Murine NEK6

TAGTGGGTGTCCGGGGCCCGGGGCGGGGGTCCGGGCAGGGGTCCGGTGCGTTTCCAGGGC
TTTCCCGAAGAAAATGCACGCAGCGTTTCCTCTGGATGACGGCCTGTGCTTTCCCTAGGA
CCCGGCAGTTAGGCTCTGGAGTGTGTTACTGATTCATCCTCTCTCTGCCTCTGTCATCGC
AGGTTTAAACTTAAACGATCGCTCCCCGAGGAGGAGGGGACCCGCCGCGCGATTGACAGT
CTATTCCTGCAGGCGCCGTCCCTCAGCCCCCACCCCACGGCCGGATTCCCGGGGCTCTGC
TCAGGGCTGAAACCGGAGAAGAAATCCTTAGATCTCTTTTCTTTACAACAAATACGAACC
TAGAAACCATCGGCATTTTGCCTTGGCATTTCATATTGGAAAGATGAAGAAGTTTTTCGA
CTCCAGGCGGGAGCAGGGGAGCTCTGGCCTGGGTTCAGGCTCCAGCGGAGGAGGGGGCAG
CAGCTCAGGCCTGGGCAGTGGTTACATCGGAAGGGTCTTCGGCATCGGGAGGCAGCAGGT
CACCGTGGACGAGGTGTTGGCGGAAGGTGGATTCGCTCTTGTCTTTCTGGTGAGAACAAG
CAATGGGGTGAAATGTGCCTTGAAACGTATGTTTGTCAACAATGAACATGACCTCCAGGT
GTGCAAGAGGGAGATCCAGATCATGAGAGACCTATCAGGGCACAAGAACATTGTGGGCTA
CATTGATTCTAGTATCAACAATGTGAGCAGCGGCGACGTCTGGGAAGTTCTCATTCTCAT
GGACTTCTGCAGAGGAGGCCAGGTGGTAAACCTGATGAACCAGCGTCTCCAGACAGGCTT

FIG. 4C

CACAGAGAATGAAGTGCTGCAGATATTCTGTGACACCTGTGAAGCTGTGGCCCGCCTGCA
TCAGTGCAAAACTCCCATTATCCACCGGGACCTGAAGGTTGAAAATATCCTCTTGCATGA
CCGAGGCCACTATGTCTTGTGTGACTTTGGAAGTGCCACCAACAAATTCCAGAATCCACA
GGCCGAGGGAGTCAATGCAGTAGAAGATGAGATTAAGAAATACACAACGCTCTCCTATCG
AGCCCCAGAAATGGTCAACTTGTACAGTGGCAAAATCATCACTACGAAGGCAGATATTTG
GGCTCTAGGCTGTTTGTTGTATAAATTATGCTACTTCACTTTGCCGTTTGGGGAGAGCCA
GGTGGCGATTTGTGACGGAAGCTTCACAATTCCTGATAACTCTCGTTATTCTCAAGATAT
GCACTGCCTTATTAGGTATATGTTGGAACCAGACCCTGACAAAAGGCCAG

FIG. 4D

Human NEK4a long C-term isoform
N-terminal=1-42 kinase=43-294 C-terminal=295-369
MAGQPGHMPHGGSSNNLCHTLGPVHPPDPQRHPNTLSFRCSLADFQIEKKIGRGQFSEVYKA
TCLLDRKTVALKKVQIFEMMDAKARQDCVKEIGLLKQLNHPNIIKYLDSFIEDNELNIVLEL
ADAGDLSQMIKYFKKQKRLIPERTVWKYFVQLCSAVEHMHSRRVMHRDIKPANVFITATGVV
KLGDLGLGRFFSSETTAAHSLVGTPYYMSPERIHENGYNFKSDIWSLGCLLYEMAALQSPFY
GDKMNLFSLCQKIEQCDYPPLPGEHYSEKLRELVSMCICPDPHQRPGSLEQLRPQGSAGSPK
GCPALQQMLKAEQLREGRWPHVTDGQIPKSFLYTVVDNLSWVNKGRWFSEPRQPPVSGL

Human NEK4b short C-term isoform
kinase (partial)=1-250 C-terminal=251-268
QIEKKIGRGQFSEVYKATCLLDRKTVALKKVQIFEMMDAKARQDCVKEIGLLKQLNHPNIIK
YLDSFIEDNELNIVLELADAGDLSQMIKYFKKQKRLIPERTVWKYFVQLCSAVEHMHSRRVM
HRDIKPANVFITATGVVKLGDLGLGRFFSSETTAAHSLVGTPYYMSPERIHENGYNFKSDIW
SLGCLLYEMAALQSPFYGDKMNLFSLCQKIEQCDYPPLPGEHYSEKLRELVSMCICPDPHQR
PDIGYVHQVAKQMHIWMSST

Human NEK5
N-terminal=1-31 kinase=32-293 C-terminal=294-302
MDEQSQGMQGPPVPQFQPQKALRPDMGYNTLANFRIEKKIGRGQFSEVYRAACLLDGVPVAL
KKVQIFDLMDAKARADCIKEIDLLKQLNHPNVIKYYASFIEDNELNIVLELADAGDLSRMIK
HFKKQKRLIPERTVWKYFVQLCSALEHMHSRRVMHRDIKPANVFITATGVVKLGDLGLGRFF
SSKTTAAHSLVGTPYYMSPERIHENGYNFKSDIWSLGCLLYEMAALQSPFYGDKMNLYSLCK
KIEQCDYPPLPSDHYSEELRQLVNMCINPDPEKRPDVTYVYDVAKRMHACTASS

Murine NEK6
N-terminal=1-22 kinase=23-302 C-terminal=missing
MKKFFDSRREQGSSGLGSGSSGGGGSSSGLGSGYIGRVFGIGRQQVTVDEVLAEGGFALVFL
VRTSNGVKCALKRMFVNNEHDLQVCKREIQIMRDLSGHKNIVGYIDSSINNVSSGDVWEVLI
LMDFCRGGQVVNLMNQRLQTGFTENEVLQIFCDTCEAVARLHQCKTPIIHRDLKVENILLHD
RGHYVLCDFGSATNKFQNPQAEGVNAVEDEIKKYTTLSYRAPEMVNLYSGKIITTKADIWAL
GCLLYKLCYFTLPFGESQVAICDGSFTIPDNSRYSQDMHCLIRYMLEPDPDKRP

FIG. 5

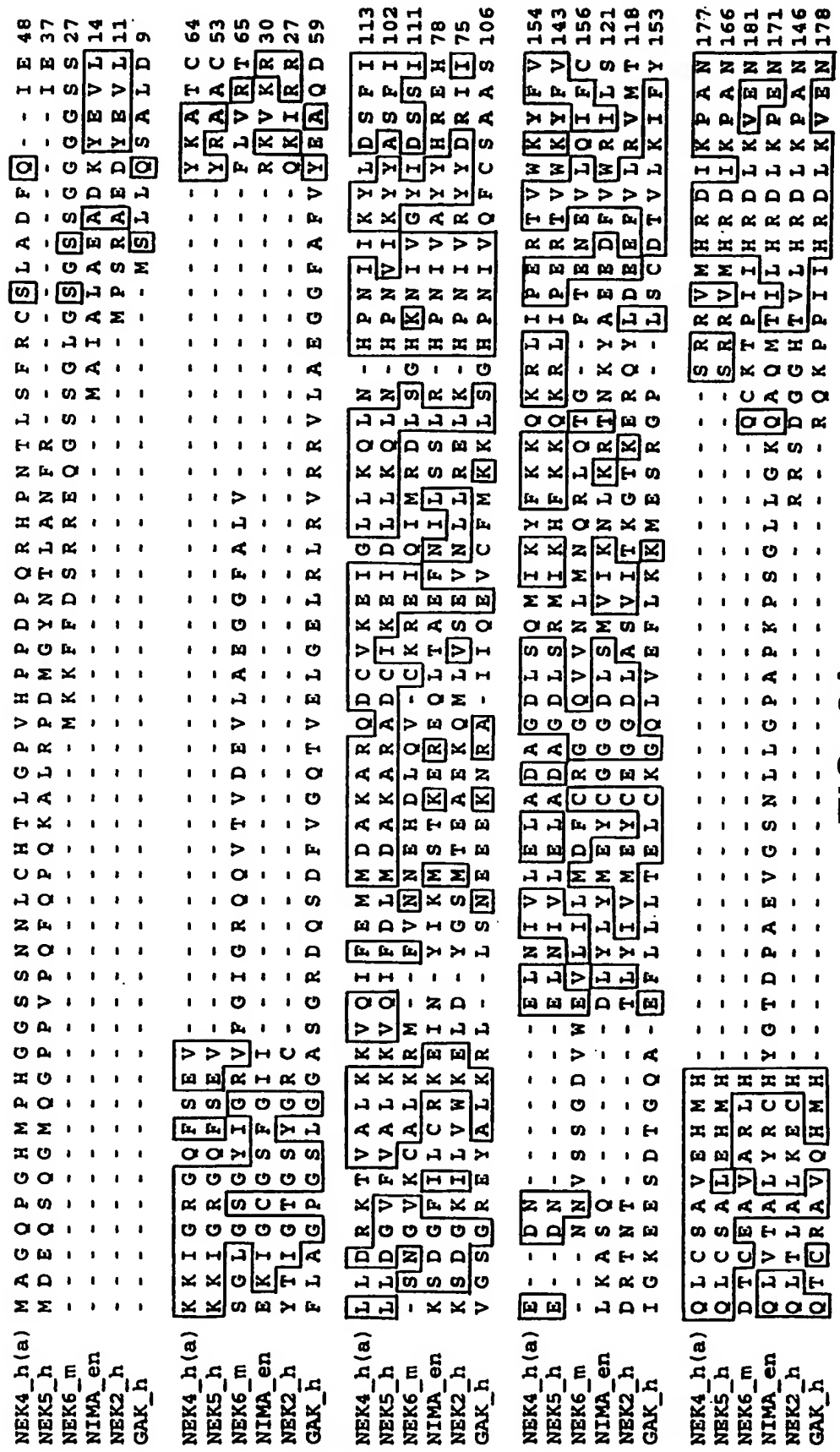

FIG. 6A

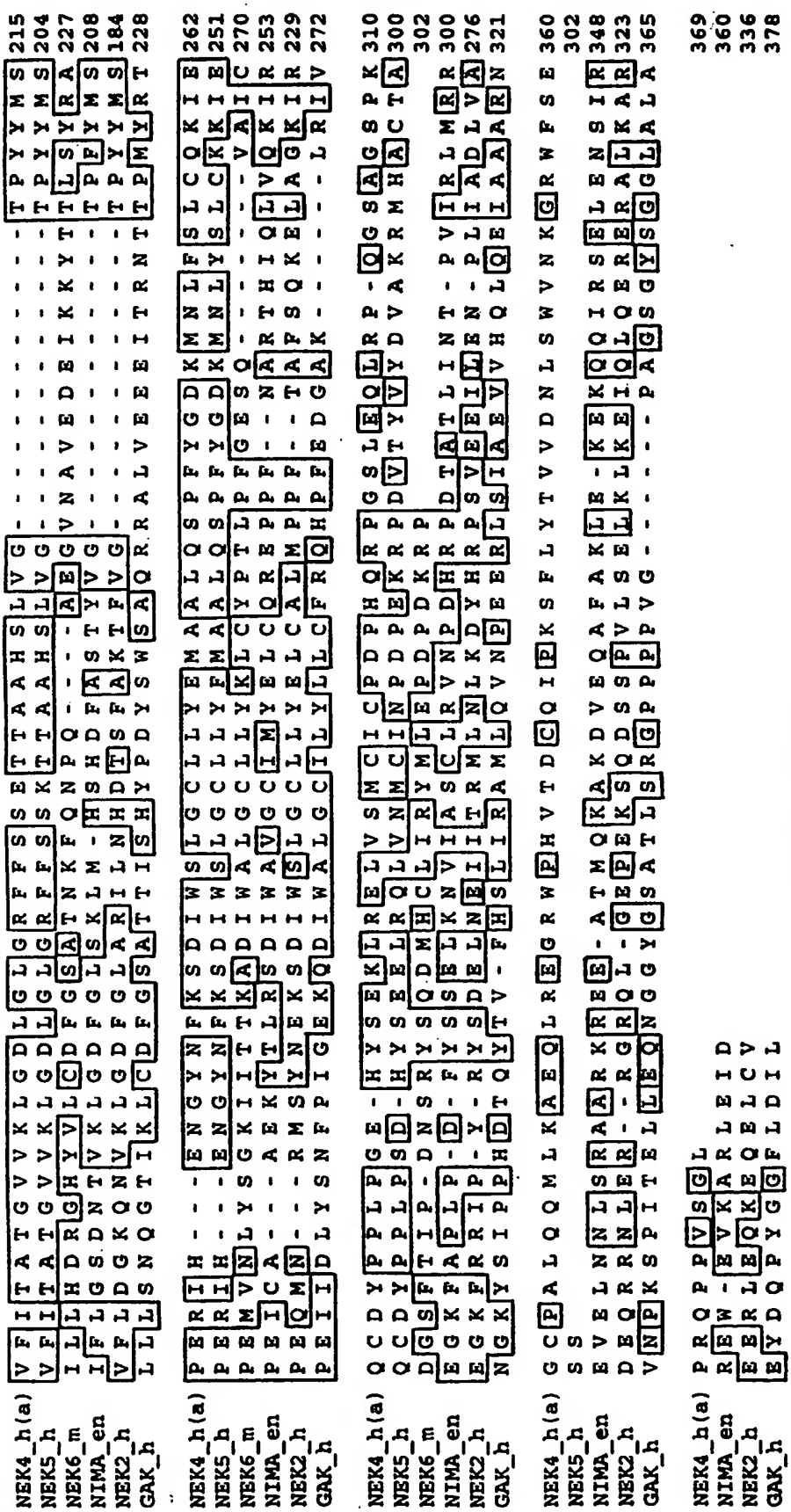

FIG. 6B

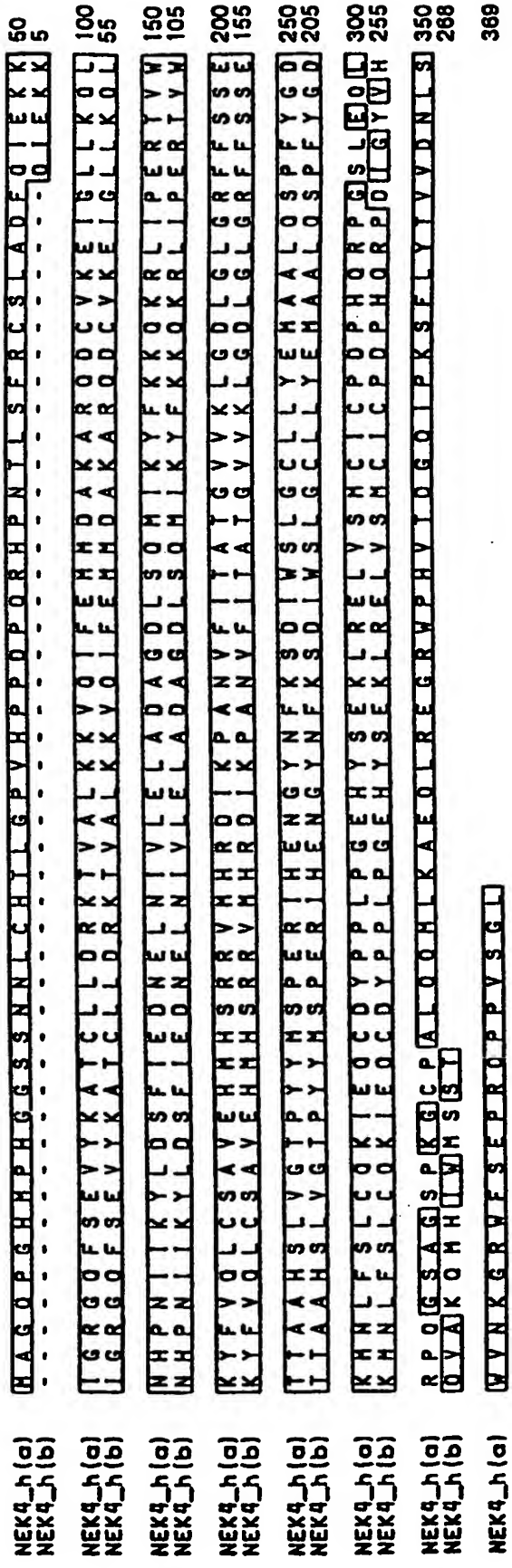
NEK4_h(a) MAGQPGHMPHGGSSNLCHTLGPVHPPDPQRHPNTLSFRCSLADFQIEKK 50
NEK4_h(b) ---------- 5
NEK4_h(a) IGRGQFSEVYKATCLLDRKTVALKKVQIFEHMDAKARQDCVKEIGLLKQL 100
NEK4_h(b) IGRGQFSEVYKATCLLDRKTVALKKVQIFEHMDAKARQDCVKEIGLLKQL 55
NEK4_h(a) NHPNIIKYLDSFIEDNELNIVLELADAGDLSQMIKYFKKQKRLIPERTVW 150
NEK4_h(b) NHPNIIKYLDSFIEDNELNIVLELADAGDLSQMIKYFKKQKRLIPERTVW 105
NEK4_h(a) KYFVQLCSAVEHMHSRRVMHRDIKPANVFITATGVVKLGDLGLGRFFSSE 200
NEK4_h(b) KYFVQLCSAVEHMHSRRVMHRDIKPANVFITATGVVKLGDLGLGRFFSSE 155
NEK4_h(a) TTAAHSLVGTPYYMSPERIHENGYNFKSDIWSLGCLLYEMAALQSPFYGD 250
NEK4_h(b) TTAAHSLVGTPYYMSPERIHENGYNFKSDIWSLGCLLYEMAALQSPFYGD 205
NEK4_h(a) KHNLFSLCQKIEQCDYPPLPGEHYSEKLRELVSMCICPDPHQRPDIGYVH 300
NEK4_h(b) KHNLFSLCQKIEQCDYPPLPGEHYSEKLRELVSMCICPDPHQRPDIGYVH 255
NEK4_h(a) RPQGSAGSPKGCPALQQHLKAEQLREGRWPHVIQGQIPKSFLYIVVDNLS 350
NEK4_h(b) QVAKQMHLWMSST 268
NEK4_h(a) WVNKGRWFSEPRQPPVSGL 369

FIG. 7

SEQUENCE LISTING

<110> Gregory D. Plowman
Ricardo Martinez
Yingfang Zhu

<120> NEK-RELATED AND BUB1-RELATED PROTEIN KINASES

<130> 243/005-PCT

<150> 60/098,265
<151> 1998-08-28

<150> 60/089,520
<151> 1998-06-16

<160> 24

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 3408
<212> DNA
<213> BUB1 human nucleotide

<400> 1

gtttgcggtt caggtttggc cgctgccggc cagcgtcctc tggccatgga caccccggaa 60
aatgtccttc agatgcttga agcccacatg cagagctaca agggcaatga ccctcttggt 120
gaatgggaaa gatacataca gtgggtagaa gagaattttc ctgagaataa agaatacttg 180
ataactttac tagaacattt aatgaaggaa tttttagata agaagaaata ccacaatgac 240
ccaagattca tcagttattg tttaaaattt gctgagtaca acagtgacct ccatcaattt 300
tttgagtttc tgtacaacca tgggattgga accctgtcat ccctctctgta cattgcctgg 360
gcggggcatc tggaagccca aggagagctg cagcatgcca gtgctgtcct tcaaagagga 420
attcaaaacc aggctgaacc cagagagttc ctgcaacaac aatacaggtt atttcagaca 480
cgcctcactg aaacccattt gccagctcaa gctagaacct cagaacctct gcataatgtt 540
caggttttaa atcaaatgat aacatcaaaa tcaaatccag gaaataacat ggcctgcatt 600
tctaagaatc agggttcaga gctttctgga gtgatatctt cagcttgtga taaagagtca 660
aatatggaac gaagagtgat cacgatttct aaatcagaat attctgtgca ctcatctttg 720
gcatccaaag ttgatgttga gcaggttgtt atgtattgca aggagaagct tattcgtggg 780
gaatcagaat tttcctttga agaattgaga gcccagaaat acaatcaacg gagaaagcat 840
gagcaatggg taaatgaaga cagacattat atgaaaagga aagaagcaaa tgcttttgaa 900
gaacagctat taaaacagaa aatggatgaa cttcataaga agttgcatca ggtggtggag 960
acatcccatg aggatctgcc cgcttcccag gaaaggtccg aggttaatcc agcacgtatg 1020
gggccaagtg taggctccca gcaggaactg agagcgccat gtcttccagt aacctatcag 1080
cagacaccag tgaacatgga aaagaaccca agagaggcac ctcctgttgt tcctcctttg 1140
gcaaatgcta tttctgcagc tttggtgtcc ccagccacca gccagagcat tgctcctcct 1200
gttcctttga aagcccagac agtaacagac tccatgtttg cagtggccag caaagatgct 1260
ggatgtgtga ataagagtac tcatgaattc aagccacaga gtggagcaga gatcaaagaa 1320
gggtgtgaaa cacataaggt tgccaacaca agttcttttc acacaactcc aaacacatca 1380
ctgggaatgg ttcaggcaac gccatccaaa gtgcagccat cacccaccgt gcacacaaaa 1440

```
gaagcattag gtttcatcat gaatatgttt caggctccta cacttcctga tatttctgat    1500
gacaaagatg aatggcaatc tctagatcaa aatgaagatg catttgaagc ccagtttcaa    1560
aaaaatgtaa ggtcatctgg ggcttgggga gtcaataaga tcatctcttc tttgtcatct    1620
gcttttcatg tgtttgaaga tggaaacaaa gaaaattatg gattaccaca gcctaaaaat    1680
aaacccacag gagccaggac ctttggagaa cgctctgtca gcagacttcc ttcaaaacca    1740
aaggaggaag tgcctcatgc tgaagagttt ttggatgact caactgtatg gggtattcgc    1800
tgcaacaaaa ccctggcacc cagtcctaag agcccaggag acttcacatc tgctgcacaa    1860
cttgcgtcta caccattcca caagcttcca gtggagtcag tgcacatttt agaagataaa    1920
gaaaatgtgg tagcaaaaca gtgtacccag gcgactttgg attcttgtga ggaaaacatg    1980
gtggtgcctt caagggatgg aaaattcagt ccaattcaag agaaaagccc aaaacaggcc    2040
ttgtcgtctc acatgtattc agcatcctta cttcgtctga gccagcctgc tgcaggtggg    2100
gtacttacct gtgaggcaga gttgggcgtt gaggcttgca gactcacaga cactgacgct    2160
gccattgcag aagatccacc agatgctatt gctgggctcc aagcagaatg gatgcagatg    2220
agttcacttg ggactgttga tgctccaaac ttcattgttg ggaacccatg ggatgataag    2280
ctgattttca aacttttatc tgggctttct aaaccagtga gttcctatcc aaatactttt    2340
gaatggcaat gtaaacttcc agccatcaag cccaagactg aatttcaatt gggttctaag    2400
ctggtctatg tccatcacct tcttggagaa ggagcctttg cccaggtgta cgaagctacc    2460
cagggagatc tgaatgatgc taaaaataaa cagaaatttg ttttaaaggt ccaaaagcct    2520
gccaacccct gggaattcta cattgggacc cagttgatgg aaagactaaa gccatctatg    2580
cagcacatgt ttatgaagtt ctattctgcc cacttattcc agaatggcag tgtattagta    2640
ggagagctct acagctatgg aacattatta aatgccatta acctctataa aaataccct     2700
gaaaaagtga tgcctcaagg tcttgtcatc tcttttgcta tgagaatgct ttacatgatt    2760
gagcaagtgc atgactgtga aatcattcat ggagacatta aaccagacaa tttcatactt    2820
ggaaacggat ttttggaaca ggatgatgaa gatgatttat ctgctggctt ggcactgatt    2880
gacctgggtc agagtataga tatgaaactt tttccaaaag gaactatatt cacagcaaag    2940
tgtgaaacat ctggttttca gtgtgttgag atgctcagca acaaaccatg gaactaccag    3000
atcgattact ttggggttgc tgcaacagta tattgcatgc tctttggcac ttacatgaaa    3060
gtgaaaaatg aaggaggaga gtgtaagcct gaaggtcttt ttagaaggct tcctcatttg    3120
gatatgtgga atgaattttt tcatgttatg ttgaatattc cagattgtca tcatcttcca    3180
tctttggatt tgttaaggca aaagctgaag aaagtatttc aacaacacta tactaacaag    3240
attagggccc tacgtaatag gctaattgta ctgctcttag aatgtaagcg ttcacgaaaa    3300
taaaatttgg atatagacag tccttaaaaa tcacactgta aatatgaatc tgctcacttt    3360
aaacctgttt ttttttcatt tattgtttat gtaaatgttt gtcaaaaa                 3408
```

<210> 2
<211> 1085
<212> PRT
<213> BUB1 aa

<400> 2

```
Met Asp Thr Pro Glu Asn Val Leu Gln Met Leu Glu Ala His Met Gln
 1               5                  10                  15

Ser Tyr Lys Gly Asn Asp Pro Leu Gly Glu Trp Glu Arg Tyr Ile Gln
            20                  25                  30

Trp Val Glu Glu Asn Phe Pro Glu Asn Lys Glu Tyr Leu Ile Thr Leu
        35                  40                  45

Leu Glu His Leu Met Lys Glu Phe Leu Asp Lys Lys Lys Tyr His Asn
    50                  55                  60
```

```
Asp Pro Arg Phe Ile Ser Tyr Cys Leu Lys Phe Ala Glu Tyr Asn Ser
65                  70                  75                  80

Asp Leu His Gln Phe Phe Glu Phe Leu Tyr Asn His Gly Ile Gly Thr
                85                  90                  95

Leu Ser Ser Pro Leu Tyr Ile Ala Trp Ala Gly His Leu Glu Ala Gln
            100                 105                 110

Gly Glu Leu Gln His Ala Ser Ala Val Leu Gln Arg Gly Ile Gln Asn
        115                 120                 125

Gln Ala Glu Pro Arg Glu Phe Leu Gln Gln Gln Tyr Arg Leu Phe Gln
    130                 135                 140

Thr Arg Leu Thr Glu Thr His Leu Pro Ala Gln Ala Arg Thr Ser Glu
145                 150                 155                 160

Pro Leu His Asn Val Gln Val Leu Asn Gln Met Ile Thr Ser Lys Ser
                165                 170                 175

Asn Pro Gly Asn Asn Met Ala Cys Ile Ser Lys Asn Gln Gly Ser Glu
            180                 185                 190

Leu Ser Gly Val Ile Ser Ser Ala Cys Asp Lys Glu Ser Asn Met Glu
        195                 200                 205

Arg Arg Val Ile Thr Ile Ser Lys Ser Glu Tyr Ser Val His Ser Ser
    210                 215                 220

Leu Ala Ser Lys Val Asp Val Glu Gln Val Val Met Tyr Cys Lys Glu
225                 230                 235                 240

Lys Leu Ile Arg Gly Glu Ser Glu Phe Ser Phe Glu Glu Leu Arg Ala
                245                 250                 255

Gln Lys Tyr Asn Gln Arg Arg Lys His Glu Gln Trp Val Asn Glu Asp
            260                 265                 270

Arg His Tyr Met Lys Arg Lys Glu Ala Asn Ala Phe Glu Glu Gln Leu
        275                 280                 285

Leu Lys Gln Lys Met Asp Glu Leu His Lys Lys Leu His Gln Val Val
    290                 295                 300

Glu Thr Ser His Glu Asp Leu Pro Ala Ser Gln Glu Arg Ser Glu Val
305                 310                 315                 320

Asn Pro Ala Arg Met Gly Pro Ser Val Gly Ser Gln Gln Glu Leu Arg
                325                 330                 335

Ala Pro Cys Leu Pro Val Thr Tyr Gln Gln Thr Pro Val Asn Met Glu
            340                 345                 350
```

```
Lys Asn Pro Arg Glu Ala Pro Pro Val Val Pro Pro Leu Ala Asn Ala
        355                 360                 365

Ile Ser Ala Ala Leu Val Ser Pro Ala Thr Ser Gln Ser Ile Ala Pro
    370                 375                 380

Pro Val Pro Leu Lys Ala Gln Thr Val Thr Asp Ser Met Phe Ala Val
385                 390                 395                 400

Ala Ser Lys Asp Ala Gly Cys Val Asn Lys Ser Thr His Glu Phe Lys
                405                 410                 415

Pro Gln Ser Gly Ala Glu Ile Lys Glu Gly Cys Glu Thr His Lys Val
            420                 425                 430

Ala Asn Thr Ser Ser Phe His Thr Thr Pro Asn Thr Ser Leu Gly Met
        435                 440                 445

Val Gln Ala Thr Pro Ser Lys Val Gln Pro Ser Pro Thr Val His Thr
    450                 455                 460

Lys Glu Ala Leu Gly Phe Ile Met Asn Met Phe Gln Ala Pro Thr Leu
465                 470                 475                 480

Pro Asp Ile Ser Asp Asp Lys Asp Glu Trp Gln Ser Leu Asp Gln Asn
                485                 490                 495

Glu Asp Ala Phe Glu Ala Gln Phe Gln Lys Asn Val Arg Ser Ser Gly
            500                 505                 510

Ala Trp Gly Val Asn Lys Ile Ile Ser Ser Leu Ser Ser Ala Phe His
        515                 520                 525

Val Phe Glu Asp Gly Asn Lys Glu Asn Tyr Gly Leu Pro Gln Pro Lys
    530                 535                 540

Asn Lys Pro Thr Gly Ala Arg Thr Phe Gly Glu Arg Ser Val Ser Arg
545                 550                 555                 560

Leu Pro Ser Lys Pro Lys Glu Glu Val Pro His Ala Glu Glu Phe Leu
                565                 570                 575

Asp Asp Ser Thr Val Trp Gly Ile Arg Cys Asn Lys Thr Leu Ala Pro
            580                 585                 590

Ser Pro Lys Ser Pro Gly Asp Phe Thr Ser Ala Ala Gln Leu Ala Ser
        595                 600                 605

Thr Pro Phe His Lys Leu Pro Val Glu Ser Val His Ile Leu Glu Asp
    610                 615                 620

Lys Glu Asn Val Val Ala Lys Gln Cys Thr Gln Ala Thr Leu Asp Ser
625                 630                 635                 640
```

```
Cys Glu Glu Asn Met Val Val Pro Ser Arg Asp Gly Lys Phe Ser Pro
                645                 650                 655

Ile Gln Glu Lys Ser Pro Lys Gln Ala Leu Ser Ser His Met Tyr Ser
            660                 665                 670

Ala Ser Leu Leu Arg Leu Ser Gln Pro Ala Ala Gly Gly Val Leu Thr
        675                 680                 685

Cys Glu Ala Glu Leu Gly Val Glu Ala Cys Arg Leu Thr Asp Thr Asp
    690                 695                 700

Ala Ala Ile Ala Glu Asp Pro Pro Asp Ala Ile Ala Gly Leu Gln Ala
705                 710                 715                 720

Glu Trp Met Gln Met Ser Ser Leu Gly Thr Val Asp Ala Pro Asn Phe
                725                 730                 735

Ile Val Gly Asn Pro Trp Asp Asp Lys Leu Ile Phe Lys Leu Leu Ser
            740                 745                 750

Gly Leu Ser Lys Pro Val Ser Ser Tyr Pro Asn Thr Phe Glu Trp Gln
        755                 760                 765

Cys Lys Leu Pro Ala Ile Lys Pro Lys Thr Glu Phe Gln Leu Gly Ser
    770                 775                 780

Lys Leu Val Tyr Val His His Leu Leu Gly Glu Gly Ala Phe Ala Gln
785                 790                 795                 800

Val Tyr Glu Ala Thr Gln Gly Asp Leu Asn Asp Ala Lys Asn Lys Gln
                805                 810                 815

Lys Phe Val Leu Lys Val Gln Lys Pro Ala Asn Pro Trp Glu Phe Tyr
            820                 825                 830

Ile Gly Thr Gln Leu Met Glu Arg Leu Lys Pro Ser Met Gln His Met
        835                 840                 845

Phe Met Lys Phe Tyr Ser Ala His Leu Phe Gln Asn Gly Ser Val Leu
    850                 855                 860

Val Gly Glu Leu Tyr Ser Tyr Gly Thr Leu Leu Asn Ala Ile Asn Leu
865                 870                 875                 880

Tyr Lys Asn Thr Pro Glu Lys Val Met Pro Gln Gly Leu Val Ile Ser
                885                 890                 895

Phe Ala Met Arg Met Leu Tyr Met Ile Glu Gln Val His Asp Cys Glu
            900                 905                 910

Ile Ile His Gly Asp Ile Lys Pro Asp Asn Phe Ile Leu Gly Asn Gly
        915                 920                 925
```

```
Phe Leu Glu Gln Asp Asp Glu Asp Asp Leu Ser Ala Gly Leu Ala Leu
    930                 935                 940

Ile Asp Leu Gly Gln Ser Ile Asp Met Lys Leu Phe Pro Lys Gly Thr
945                 950                 955                 960

Ile Phe Thr Ala Lys Cys Glu Thr Ser Gly Phe Gln Cys Val Glu Met
                965                 970                 975

Leu Ser Asn Lys Pro Trp Asn Tyr Gln Ile Asp Tyr Phe Gly Val Ala
            980                 985                 990

Ala Thr Val Tyr Cys Met Leu Phe Gly Thr Tyr Met Lys Val Lys Asn
        995                1000                1005

Glu Gly Gly Glu Cys Lys Pro Glu Gly Leu Phe Arg Arg Leu Pro His
   1010                1015                1020

Leu Asp Met Trp Asn Glu Phe Phe His Val Met Leu Asn Ile Pro Asp
1025                1030                1035                1040

Cys His His Leu Pro Ser Leu Asp Leu Leu Arg Gln Lys Leu Lys Lys
                1045                1050                1055

Val Phe Gln Gln His Tyr Thr Asn Lys Ile Arg Ala Leu Arg Asn Arg
            1060                1065                1070

Leu Ile Val Leu Leu Leu Glu Cys Lys Arg Ser Arg Lys
        1075                1080                1085
```

<210> 3
<211> 3854
<212> DNA
<213> Human NEK5 nucleotide

<400> 3

```
ctgagttcta aagttcctgt tgcttcagac aatggatgag caatcacaag gaatgcaagg      60
gccacctgtt cctcagttcc aaccacagaa ggccttacga ccggatatgg gctataatac     120
attagccaac tttcgaatag aaaagaaaat tggtcgcgga caatttagtg aagtttatag     180
agcagcctgt ctcttggatg gagtaccagt agctttaaaa aaagtgcaga tatttgattt     240
aatggatgcc aaagcacgtg ctgattgcat caaagaaata gatcttctta agcaactcaa     300
ccatccaaat gtaataaaat attatgcatc attcattgaa gataatgaac taaacatagt     360
tttggaacta gcagatgctg gcgacctatc cagaatgatc aagcatttta agaagcaaaa     420
gaggctaatt cctgaaagaa ctgtttggaa gtattttgtt cagctttgca gtgcattgga     480
acacatgcat tctcgaagag tcatgcatag agatataaaa ccagctaatg tgttcattac     540
agccactggg gtggtaaaac ttggagatct tgggcttggc cggtttttca gctcaaaaac     600
cacagctgca cattctttag ttggtacgcc ttattacatg tctccagaga gaatacatga     660
aaatggatac aacttcaaat ctgacatctg gtctcttggc tgtctactat atgagatggc     720
tgcattacaa agtcctttct atggtgacaa aatgaattta tactcactgt gtaagaagat     780
agaacagtgt gactacccac ctcttccttc agatcactat tcagaagaac tccgacagtt     840
```

agttaatatg tgcatcaacc cagatccaga gaagcgacca gacgtcacct atgtttatga 900
cgtagcaaag aggatgcatg catgcactgc aagcagctaa acatgcaaga tcatgaagag 960
tgtaaccaaa gtaattgaaa gtattttgtg caagtcatac ctccccattt atgtctggtg 1020
ttaagattaa tatttcagag ctagtgtgct ttgaatcctt aaccagtttt catataagct 1080
tcattttgta ccagtcacct aaatcacctc cttgcaaccc ccaaatgact ttggaataac 1140
tgaattgcat gttaggagag aaaatgaaac atgatggttt tgaatggcta aaggtttata 1200
gaatttctta cagttttctg ctgataaatt gtgtttagat agactgtcag tgccaaatat 1260
tgaaggtgca gcttggcaca catcagaata gactcatacc tgagaaaaag tatctgaaca 1320
tgtgacttgt ttctttttta gtaatttatg gacattgaga tgaacacaat tgtgaacttt 1380
tgtgaagatt ttatttttaa acgtttgaag tactagtttt agttcttagc agagtagttt 1440
tcaaatatga ttcttatgat aaatgtagac acaaactatt tgagaaacat ttagaactct 1500
tagcttatac attcaaaatg taactattaa atgtgaagat ttggggacaa aatgtgagtc 1560
agacactgaa gagtttttg ttttgtttta atatttttga tattctcttt gcattgaaat 1620
ggtataaatg aatccattta aaaagtggtt aaggatttgt ttagctggtg tgataataat 1680
ttttaaagtt gcacattgcc caaggctttt tttgtgtgtt tttattgttg tttgtacatt 1740
tgaaaaatat tctttgaata accttgcagt actatatttc aatttcttta taaatttaag 1800
tgcattttaa ctcataattg tacactataa tataagccta agtttttatt cataagtttt 1860
attgaagttc tgatcggtcc ccttcagaaa tttttttata ttattcttca agttactttc 1920
ttatttatat tgtatgtgca ttttatccat taatgtttca tactttctga gagtataata 1980
ccctttaaa agatatttgg tataccaata cttttcctgg attgaaaact ttttttaaac 2040
tttttaaaat ttgggccact ctgtatgcat atgtttggtc ttgttaaaga ggaagaaagg 2100
atgtgtgtta tactgtacct gtgaatgttg atacagttac aatttatttg acaaggttgt 2160
aattctagaa tatgcttaat aaaatgaaaa ctggccatga ctacagccag aactgttatg 2220
agattaacat ttctattgag aagctttga gtaaagtact gtatttgttc atgaagatga 2280
ctgagatggt aacacttcgt gtagcttaag gaaatgggca gaatttcgta aatgctgttg 2340
tgcagatgtg ttttccctga atgctttcgt attagtggcg accagtttct cacagaattg 2400
tgaagcctga aggccaagag gaagtcactg ttaaaggact ctgtgccatc ttacaacctt 2460
ggatgaatta tcctgccaac gtgaaaacct catgttcaaa gaacacttcc ctttagccga 2520
tgtaactgct ggttttgttt ttcatatgtg tttttcttac actcatttga atgctttcaa 2580
gcatttgtaa acttaaaaaa tgtataaagg gcaaaaagtc tgaacccttg ttttctgaaa 2640
tctaatcagt tatgtatggt ttctgaaggg taattttatt ttggaatagg taaaggaaac 2700
ctgttttgtt tgtttttcct gagggctaga tgcatttttt ttctcacact cttaatgact 2760
tttaacattt atactgagca tccatagata tattcctaga agtatgagaa gaattattct 2820
tattgaccat taatgtcatg ttcattttaa tgtaatataa ttgagatgaa atgttctctg 2880
gttggaacag atactctctt ttttttcttg caatctttaa gaatacatag atctaaaatt 2940
cattagcttg acccctcaaa gtaactttta agtaaagatt aaagctttc ttctcagtga 3000
atatatctgc tagaaggaaa tagctgggaa gaatttaatg atcagggaaa ttcattattt 3060
ctatatgtgg aaacttttg cttcgaatat tgtatctttt taaatctaaa tgttcatatt 3120
tttcctgaag aaaccactgt gtaaaaatca aattttaatt ttgaatggaa taatttcaaa 3180
gaactatgaa gatgatttga agctctaatt tatatagtca cctataaaat gttctttata 3240
tgtgttcata agtaaatttt atattgatta agttaaactt ttgaattgat ttgaggagca 3300
gtaaaatgaa agctatatct attctaaacc ttatttagac attggtacca gttacccagg 3360
tgaaaatatg gagtaacttt gttttgtatg gtaaggttta ggaatggtgg atgaagggta 3420
tctctatata aataaagtgc tcaacaatgt gcaatgattg taaatttagt aagatattac 3480
agccatttca tgaatgcttt accattcaac atagtatcta ttacaaaaca cctttcttgt 3540
atccatatac ttcaggtgtt gctgttaaca tttactatga tatttatttt aaccaaaatg 3600
ttactcacat taaatgttta ttctttaaaa tgaatgtatt atgtttttaa cccacaaatg 3660
catacttacc ctgtgcctca tatttcaata gtactgtaat atggacatct tttgtgaaat 3720
acttttattt tgttatgctt taaatataca tacaaaaaga tttctgttat tagctttgaa 3780
aattgtataa tatcctaata taaacaaaaa tataaaaata aaaatgaata cagtaaaaaa 3840
aaaaaaaaaa aaaa 3854

<210> 4
<211> 302
<212> PRT
<213> Human NEK5 aa

<400> 4

```
Met Asp Glu Gln Ser Gln Gly Met Gln Gly Pro Pro Val Pro Gln Phe
 1               5                  10                  15

Gln Pro Gln Lys Ala Leu Arg Pro Asp Met Gly Tyr Asn Thr Leu Ala
            20                  25                  30

Asn Phe Arg Ile Glu Lys Lys Ile Gly Arg Gly Gln Phe Ser Glu Val
        35                  40                  45

Tyr Arg Ala Ala Cys Leu Leu Asp Gly Val Pro Val Ala Leu Lys Lys
    50                  55                  60

Val Gln Ile Phe Asp Leu Met Asp Ala Lys Ala Arg Ala Asp Cys Ile
65                  70                  75                  80

Lys Glu Ile Asp Leu Leu Lys Gln Leu Asn His Pro Asn Val Ile Lys
                85                  90                  95

Tyr Tyr Ala Ser Phe Ile Glu Asp Asn Glu Leu Asn Ile Val Leu Glu
            100                 105                 110

Leu Ala Asp Ala Gly Asp Leu Ser Arg Met Ile Lys His Phe Lys Lys
        115                 120                 125

Gln Lys Arg Leu Ile Pro Glu Arg Thr Val Trp Lys Tyr Phe Val Gln
    130                 135                 140

Leu Cys Ser Ala Leu Glu His Met His Ser Arg Arg Val Met His Arg
145                 150                 155                 160

Asp Ile Lys Pro Ala Asn Val Phe Ile Thr Ala Thr Gly Val Val Lys
                165                 170                 175

Leu Gly Asp Leu Gly Leu Gly Arg Phe Phe Ser Ser Lys Thr Thr Ala
            180                 185                 190

Ala His Ser Leu Val Gly Thr Pro Tyr Tyr Met Ser Pro Glu Arg Ile
        195                 200                 205

His Glu Asn Gly Tyr Asn Phe Lys Ser Asp Ile Trp Ser Leu Gly Cys
    210                 215                 220

Leu Leu Tyr Glu Met Ala Ala Leu Gln Ser Pro Phe Tyr Gly Asp Lys
225                 230                 235                 240

Met Asn Leu Tyr Ser Leu Cys Lys Lys Ile Glu Gln Cys Asp Tyr Pro
                245                 250                 255
```

```
Pro Leu Pro Ser Asp His Tyr Ser Glu Glu Leu Arg Gln Leu Val Asn
            260                 265                 270

Met Cys Ile Asn Pro Asp Pro Glu Lys Arg Pro Asp Val Thr Tyr Val
        275                 280                 285

Tyr Asp Val Ala Lys Arg Met His Ala Cys Thr Ala Ser Ser
    290                 295                 300
```

<210> 5
<211> 1310
<212> DNA
<213> Murine NEK6 nucleotide

<400> 5

```
tagtgggtgt ccggggcccg gggcgggggt ccgggcaggg gtccggtgcg tttccagggc      60
tttcccgaag aaaatgcacg cagcgtttcc tctggatgac ggcctgtgct ttccctagga     120
cccggcagtt aggctctgga gtgtgttact gattcatcct ctctctgcct ctgtcatcgc     180
aggtttaaac ttaaacgatc gctccccgag gaggagggga cccgccgcgc gattgacagt     240
ctattcctgc aggcgccgtc cctcagcccc caccccacgg ccggattccc ggggctctgc     300
tcagggctga aaccggagaa gaaatcctta gatctctttt ctttacaaca aatacgaacc     360
tagaaaccat cggcattttg ccttggcatt tcatattgga aagatgaaga agtttttcga     420
ctccaggcgg gagcagggga gctctggcct gggttcaggc tccagcggag gagggggcag     480
cagctcaggc ctgggcagtg gttacatcgg aagggtcttc ggcatcggga ggcagcaggt     540
caccgtggac gaggtgttgg cggaaggtgg attcgctctt gtctttctgg tgagaacaag     600
caatggggtg aaatgtgcct tgaaacgtat gtttgtcaac aatgaacatg acctccaggt     660
gtgcaagagg gagatccaga tcatgagaga cctatcaggg cacaagaaca ttgtgggcta     720
cattgattct agtatcaaca atgtgagcag cggcgacgtc tgggaagttc tcattctcat     780
ggacttctgc agaggaggcc aggtggtaaa cctgatgaac cagcgtctcc agacaggctt     840
cacagagaat gaagtgctgc agatattctg tgacacctgt gaagctgtgg cccgcctgca     900
tcagtgcaaa actcccatta tccaccggga cctgaaggtt gaaaatatcc tcttgcatga     960
ccgaggccac tatgtcttgt gtgactttgg aagtgccacc aacaaattcc agaatccaca    1020
ggccgaggga gtcaatgcag tagaagatga gattaagaaa tacacaacgc tctcctatcg    1080
agccccagaa atggtcaact tgtacagtgg caaaatcatc actacgaagg cagatatttg    1140
ggctctaggc tgtttgttgt ataaattatg ctacttcact ttgccgtttg gggagagcca    1200
ggtggcgatt tgtgacggaa gcttcacaat tcctgataac tctcgttatt ctcaagatat    1260
gcactgcctt attaggtata tgttggaacc agaccctgac aaaaggccag               1310
```

<210> 6
<211> 302
<212> PRT
<213> Murine NEK6 aa

<400> 6

```
Met Lys Lys Phe Phe Asp Ser Arg Arg Glu Gln Gly Ser Ser Gly Leu
 1               5                  10                  15
```

```
Gly Ser Gly Ser Ser Gly Gly Gly Gly Ser Ser Ser Gly Leu Gly Ser
            20                  25                  30

Gly Tyr Ile Gly Arg Val Phe Gly Ile Gly Arg Gln Gln Val Thr Val
        35                  40                  45

Asp Glu Val Leu Ala Glu Gly Gly Phe Ala Leu Val Phe Leu Val Arg
    50                  55                  60

Thr Ser Asn Gly Val Lys Cys Ala Leu Lys Arg Met Phe Val Asn Asn
65                  70                  75                  80

Glu His Asp Leu Gln Val Cys Lys Arg Glu Ile Gln Ile Met Arg Asp
                85                  90                  95

Leu Ser Gly His Lys Asn Ile Val Gly Tyr Ile Asp Ser Ser Ile Asn
            100                 105                 110

Asn Val Ser Ser Gly Asp Val Trp Glu Val Leu Ile Leu Met Asp Phe
        115                 120                 125

Cys Arg Gly Gly Gln Val Val Asn Leu Met Asn Gln Arg Leu Gln Thr
    130                 135                 140

Gly Phe Thr Glu Asn Glu Val Leu Gln Ile Phe Cys Asp Thr Cys Glu
145                 150                 155                 160

Ala Val Ala Arg Leu His Gln Cys Lys Thr Pro Ile Ile His Arg Asp
                165                 170                 175

Leu Lys Val Glu Asn Ile Leu Leu His Asp Arg Gly His Tyr Val Leu
            180                 185                 190

Cys Asp Phe Gly Ser Ala Thr Asn Lys Phe Gln Asn Pro Gln Ala Glu
        195                 200                 205

Gly Val Asn Ala Val Glu Asp Glu Ile Lys Lys Tyr Thr Thr Leu Ser
    210                 215                 220

Tyr Arg Ala Pro Glu Met Val Asn Leu Tyr Ser Gly Lys Ile Ile Thr
225                 230                 235                 240

Thr Lys Ala Asp Ile Trp Ala Leu Gly Cys Leu Leu Tyr Lys Leu Cys
                245                 250                 255

Tyr Phe Thr Leu Pro Phe Gly Glu Ser Gln Val Ala Ile Cys Asp Gly
            260                 265                 270

Ser Phe Thr Ile Pro Asp Asn Ser Arg Tyr Ser Gln Asp Met His Cys
        275                 280                 285

Leu Ile Arg Tyr Met Leu Glu Pro Asp Pro Asp Lys Arg Pro
    290                 295                 300
```

<210> 7
<211> 1449
<212> DNA
<213> Human NEK4a nucleotide

<400> 7

tgcgccgcaa actcgtgtgg gacgcaccgc tccagccgcc cgcgggccag cgcaccggtc 60
ccccagcggc agccgagccc gcccgcgcgc cgttcgtgcc ctcgtgaggc tggcatgcag 120
gatggcagga cagcccggcc acatgcccca tggagggagt tccaacaacc tctgccacac 180
cctggggcct gtgcatcctc ctgacccaca gaggcatccc aacacgctgt cttttcgctg 240
ctcgctggcg gacttccaga tcgaaaagaa gataggccga ggacagttca gcgaggtgta 300
caaggccacc tgcctgctgg acaggaagac agtggctctg aagaaggtgc agatctttga 360
gatgatggac gccaaggcga ggcaggactg tgtcaaggag atcggcctct tgaagcaact 420
gaaccaccca aatatcatca agtatttgga ctcgtttatc gaagacaacg agctgaacat 480
tgtgctggag ttggctgacg caggggacct ctcgcagatg atcaagtact ttaagaagca 540
gaagcggctc atcccggaga ggacagtatg gaagtacttt gtgcagctgt gcagcgccgt 600
ggagcacatg cattcacgcc gggtgatgca ccgagacatc aagcctgcca acgtgttcat 660
cacagccacg ggcgtcgtga agctcggtga ccttggtctg ggccgcttct tcagctctga 720
gaccaccgca gcccactccc tagtggggac gccctactac atgtcaccgg agaggatcca 780
tgagaacggc tacaacttca agtccgacat ctggtccttg ggctgtctgc tgtacgagat 840
ggcagccctc cagagccgct tctatggaga taagatgaat ctcttctccc tgtgccagaa 900
gatcgagcag tgtgactacc ccccactccc cggggagcac tactccgaga agttacgaga 960
actggtcagc atgtgcatct gccctgaccc ccaccagaga cctggtagcc tagaacagct 1020
aagaccacag ggttcagcag gttccccaaa aggctgccca gccttacagc agatgctgaa 1080
ggcagagcag ctgagggagg ggcgctggcc acatgtcact gatggtcaga ttccaaagtc 1140
ctttctttat actgttgtgg acaatctcag ctgggtcaat aagggcaggt ggttcagcga 1200
gccacggcag ccccctgtat ctggattgta atgtgaatct ttagggtaat tcctccagtg 1260
acctgtcaag gcttatgcta acaggagact tgcaggagac cgtgtgattt gtgtagtgag 1320
cctttgaaaa tggttagtac cgggttcagt ttagttcttg gtatcttttc aatcaagctg 1380
tgtgcttaat ttactctgtt gtaaagggat aaagtggaaa tcatttttaa aaaaaaaaaa 1440
aaaaaaaaa 1449

<210> 8
<211> 369
<212> PRT
<213> Human NEK4a aa

<400> 8

Met Ala Gly Gln Pro Gly His Met Pro His Gly Gly Ser Ser Asn Asn
1 5 10 15

Leu Cys His Thr Leu Gly Pro Val His Pro Pro Asp Pro Gln Arg His
20 25 30

Pro Asn Thr Leu Ser Phe Arg Cys Ser Leu Ala Asp Phe Gln Ile Glu
35 40 45

Lys Lys Ile Gly Arg Gly Gln Phe Ser Glu Val Tyr Lys Ala Thr Cys
50 55 60

```
Leu Leu Asp Arg Lys Thr Val Ala Leu Lys Lys Val Gln Ile Phe Glu
65                  70                  75                  80

Met Met Asp Ala Lys Ala Arg Gln Asp Cys Val Lys Glu Ile Gly Leu
                85                  90                  95

Leu Lys Gln Leu Asn His Pro Asn Ile Ile Lys Tyr Leu Asp Ser Phe
            100                 105                 110

Ile Glu Asp Asn Glu Leu Asn Ile Val Leu Glu Leu Ala Asp Ala Gly
        115                 120                 125

Asp Leu Ser Gln Met Ile Lys Tyr Phe Lys Lys Gln Lys Arg Leu Ile
    130                 135                 140

Pro Glu Arg Thr Val Trp Lys Tyr Phe Val Gln Leu Cys Ser Ala Val
145                 150                 155                 160

Glu His Met His Ser Arg Arg Val Met His Arg Asp Ile Lys Pro Ala
                165                 170                 175

Asn Val Phe Ile Thr Ala Thr Gly Val Val Lys Leu Gly Asp Leu Gly
            180                 185                 190

Leu Gly Arg Phe Phe Ser Ser Glu Thr Thr Ala Ala His Ser Leu Val
        195                 200                 205

Gly Thr Pro Tyr Tyr Met Ser Pro Glu Arg Ile His Glu Asn Gly Tyr
    210                 215                 220

Asn Phe Lys Ser Asp Ile Trp Ser Leu Gly Cys Leu Leu Tyr Glu Met
225                 230                 235                 240

Ala Ala Leu Gln Ser Pro Phe Tyr Gly Asp Lys Met Asn Leu Phe Ser
                245                 250                 255

Leu Cys Gln Lys Ile Glu Gln Cys Asp Tyr Pro Pro Leu Pro Gly Glu
            260                 265                 270

His Tyr Ser Glu Lys Leu Arg Glu Leu Val Ser Met Cys Ile Cys Pro
        275                 280                 285

Asp Pro His Gln Arg Pro Gly Ser Leu Glu Gln Leu Arg Pro Gln Gly
    290                 295                 300

Ser Ala Gly Ser Pro Lys Gly Cys Pro Ala Leu Gln Gln Met Leu Lys
305                 310                 315                 320

Ala Glu Gln Leu Arg Glu Gly Arg Trp Pro His Val Thr Asp Gly Gln
                325                 330                 335

Ile Pro Lys Ser Phe Leu Tyr Thr Val Val Asp Asn Leu Ser Trp Val
            340                 345                 350
```

```
Asn Lys Gly Arg Trp Phe Ser Glu Pro Arg Gln Pro Pro Val Ser Gly
        355                 360                 365

Leu
```

<210> 9
<211> 1346
<212> DNA
<213> Human NEK4b nucleotide

<400> 9

```
tccagatcga aaagaagata ggccgaggac agttcagcga ggtgtacaag gccacctgcc      60
tgctggacag gaagacagtg gctctgaaga aggtgcagat ctttgagatg atggacgcca     120
aggcgaggca ggactgtgtc aaggagatcg gcctcttgaa gcaactgaac cacccaaata     180
tcatcaagta tttggactcg tttatcgaag acaacgagct gaacattgtg ctggagttgg     240
ctgacgcagg ggacctctcg cagatgatca agtactttaa gaagcagaag cggctcatcc     300
cggagaggac agtatggaag tactttgtgc agctgtgcag cgccgtggag cacatgcatt     360
cacgccgggt gatgcaccga gacatcaagc ctgccaacgt gttcatcaca gccacgggcg     420
tcgtgaagct cggtgacctt ggtctgggcc gcttcttcag ctctgagacc accgcagccc     480
actccctagt ggggacgccc tactacatgt caccggagag gatccatgag aacggctaca     540
acttcaagtc cgacatctgg tccttgggct gtctgctgta cgagatggca gccctccaga     600
gccccttcta tggagataag atgaatctct tctccctgtg ccagaagatc gagcagtgtg     660
actacccccc actccccggg gagcactact ccgagaagtt acgagaactg gtcagcatgt     720
gcatctgccc tgacccccac cagagacctg acatcggata cgtgcaccag gtggccaagc     780
agatgcacat ctggatgtcc agcacctgag cgtggatgca ccgtgcctta tcaaagccag     840
caccactttg ccttacttga gtcgtcttct cttcgagtgg ccacctggta gcctagaaca     900
gctaagacca cagggttcag caggttcccc aaaaggctgc ccagccttac agcagatgct     960
gaaggcagag cagctgaggg aggggcgctg gccacatgtc actgatggtc agattccaaa    1020
gtcctttctt tatactgttg tggacaatct cagctgggtc aataagggca ggtggttcag    1080
cgagccacgg cagccccctg tatctggatt gtaatgtgaa tctttagggt aattcctcca    1140
gtgacctgtc aaggcttatg ctaacaggag acttgcagga gaccgtgtga tttgtgtagt    1200
gagcctttga aaatggttag taccgggttc agtttagttc ttagtatctt ttcaatcaag    1260
ctgtgtgctt aatttactct gttgtaaagg gataaagtgg aaatcaaaaa aaaaaaaaaa    1320
aaaaaaaaaa aaaaaaaaaa aaaaaa                                        1346
```

<210> 10
<211> 268
<212> PRT
<213> Human NEK4b aa

<400> 10

```
Gln Ile Glu Lys Lys Ile Gly Arg Gly Gln Phe Ser Glu Val Tyr Lys
 1               5                  10                  15

Ala Thr Cys Leu Leu Asp Arg Lys Thr Val Ala Leu Lys Lys Val Gln
            20                  25                  30
```

Ile Phe Glu Met Met Asp Ala Lys Ala Arg Gln Asp Cys Val Lys Glu
35 40 45

Ile Gly Leu Leu Lys Gln Leu Asn His Pro Asn Ile Ile Lys Tyr Leu
50 55 60

Asp Ser Phe Ile Glu Asp Asn Glu Leu Asn Ile Val Leu Glu Leu Ala
65 70 75 80

Asp Ala Gly Asp Leu Ser Gln Met Ile Lys Tyr Phe Lys Lys Gln Lys
85 90 95

Arg Leu Ile Pro Glu Arg Thr Val Trp Lys Tyr Phe Val Gln Leu Cys
100 105 110

Ser Ala Val Glu His Met His Ser Arg Arg Val Met His Arg Asp Ile
115 120 125

Lys Pro Ala Asn Val Phe Ile Thr Ala Thr Gly Val Val Lys Leu Gly
130 135 140

Asp Leu Gly Leu Gly Arg Phe Phe Ser Ser Glu Thr Thr Ala Ala His
145 150 155 160

Ser Leu Val Gly Thr Pro Tyr Tyr Met Ser Pro Glu Arg Ile His Glu
165 170 175

Asn Gly Tyr Asn Phe Lys Ser Asp Ile Trp Ser Leu Gly Cys Leu Leu
180 185 190

Tyr Glu Met Ala Ala Leu Gln Ser Pro Phe Tyr Gly Asp Lys Met Asn
195 200 205

Leu Phe Ser Leu Cys Gln Lys Ile Glu Gln Cys Asp Tyr Pro Pro Leu
210 215 220

Pro Gly Glu His Tyr Ser Glu Lys Leu Arg Glu Leu Val Ser Met Cys
225 230 235 240

Ile Cys Pro Asp Pro His Gln Arg Pro Asp Ile Gly Tyr Val His Gln
245 250 255

Val Ala Lys Gln Met His Ile Trp Met Ser Ser Thr
260 265

<210> 11
<211> 360
<212> PRT
<213> NIMA emericella nidulans

<400> 11

Met Ala Ile Ala Leu Ala Glu Ala Asp Lys Tyr Glu Val Leu Glu Lys
1 5 10 15

Ile Gly Cys Gly Ser Phe Gly Ile Ile Arg Lys Val Lys Arg Lys Ser
20 25 30

Asp Gly Phe Ile Leu Cys Arg Lys Glu Ile Asn Tyr Ile Lys Met Ser
35 40 45

Thr Lys Glu Arg Glu Gln Leu Thr Ala Glu Phe Asn Ile Leu Ser Ser
50 55 60

Leu Arg His Pro Asn Ile Val Ala Tyr Tyr His Arg Glu His Leu Lys
65 70 75 80

Ala Ser Gln Asp Leu Tyr Leu Tyr Met Glu Tyr Cys Gly Gly Gly Asp
85 90 95

Leu Ser Met Val Ile Lys Asn Leu Lys Arg Thr Asn Lys Tyr Ala Glu
100 105 110

Glu Asp Phe Val Trp Arg Ile Leu Ser Gln Leu Val Thr Ala Leu Tyr
115 120 125

Arg Cys His Tyr Gly Thr Asp Pro Ala Glu Val Gly Ser Asn Leu Leu
130 135 140

Gly Pro Ala Pro Lys Pro Ser Gly Leu Lys Gly Lys Gln Ala Gln Met
145 150 155 160

Thr Ile Leu His Arg Asp Leu Lys Pro Glu Asn Ile Phe Leu Gly Ser
165 170 175

Asp Asn Thr Val Lys Leu Gly Asp Phe Gly Leu Ser Lys Leu Met His
180 185 190

Ser His Asp Phe Ala Ser Thr Tyr Val Gly Thr Pro Phe Tyr Met Ser
195 200 205

Pro Glu Ile Cys Ala Ala Glu Lys Tyr Thr Leu Arg Ser Asp Ile Trp
210 215 220

Ala Val Gly Cys Ile Met Tyr Glu Leu Cys Gln Arg Glu Pro Pro Phe
225 230 235 240

Asn Ala Arg Thr His Ile Gln Leu Val Gln Lys Ile Arg Glu Gly Lys
245 250 255

Phe Ala Pro Leu Pro Asp Phe Tyr Ser Ser Glu Leu Lys Asn Val Ile
260 265 270

Ala Ser Cys Leu Arg Val Asn Pro Asp His Arg Pro Asp Thr Ala Thr
275 280 285

Leu Ile Asn Thr Pro Val Ile Arg Leu Met Arg Arg Glu Val Glu Leu
290 295 300

Asn Asn Leu Ser Arg Ala Ala Arg Lys Arg Glu Glu Ala Thr Met Gln
305 310 315 320

Lys Ala Lys Asp Val Glu Gln Ala Phe Ala Lys Leu Glu Lys Glu Lys
325 330 335

Gln Gln Ile Arg Ser Glu Leu Glu Asn Ser Ile Arg Arg Glu Trp Glu
340 345 350

Val Lys Ala Arg Leu Glu Ile Asp
355 360

<210> 12
<211> 336
<212> PRT
<213> NEK2 human aa

<400> 12

Met Pro Ser Arg Ala Glu Asp Tyr Glu Val Leu Tyr Thr Ile Gly Thr
1 5 10 15

Gly Ser Tyr Gly Arg Cys Gln Lys Ile Arg Arg Lys Ser Asp Gly Lys
20 25 30

Ile Leu Val Trp Lys Glu Leu Asp Tyr Gly Ser Met Thr Glu Ala Glu
35 40 45

Lys Gln Met Leu Val Ser Glu Val Asn Leu Leu Arg Glu Leu Lys His
50 55 60

Pro Asn Ile Val Arg Tyr Tyr Asp Arg Ile Ile Asp Arg Thr Asn Thr
65 70 75 80

Thr Leu Tyr Ile Val Met Glu Tyr Cys Glu Gly Gly Asp Leu Ala Ser
85 90 95

Val Ile Thr Lys Gly Thr Lys Glu Arg Gln Tyr Leu Asp Glu Glu Phe
100 105 110

Val Leu Arg Val Met Thr Gln Leu Thr Leu Ala Leu Lys Glu Cys His
115 120 125

Arg Arg Ser Asp Gly Gly His Thr Val Leu His Arg Asp Leu Lys Pro
130 135 140

```
Ala Asn Val Phe Leu Asp Gly Lys Gln Asn Val Lys Leu Gly Asp Phe
145                 150                 155                 160

Gly Leu Ala Arg Ile Leu Asn His Asp Thr Ser Phe Ala Lys Thr Phe
                165                 170                 175

Val Gly Thr Pro Tyr Tyr Met Ser Pro Glu Gln Met Asn Arg Met Ser
            180                 185                 190

Tyr Asn Glu Lys Ser Asp Ile Trp Ser Leu Gly Cys Leu Leu Tyr Glu
        195                 200                 205

Leu Cys Ala Leu Met Pro Pro Phe Thr Ala Phe Ser Gln Lys Glu Leu
    210                 215                 220

Ala Gly Lys Ile Arg Glu Gly Lys Phe Arg Arg Ile Pro Tyr Arg Tyr
225                 230                 235                 240

Ser Asp Glu Leu Asn Glu Ile Ile Thr Arg Met Leu Asn Leu Lys Asp
                245                 250                 255

Tyr His Arg Pro Ser Val Glu Glu Ile Leu Glu Asn Pro Leu Ile Ala
            260                 265                 270

Asp Leu Val Ala Asp Glu Gln Arg Arg Asn Leu Glu Arg Arg Gly Arg
        275                 280                 285

Gln Leu Gly Glu Pro Glu Lys Ser Gln Asp Ser Ser Pro Val Leu Ser
    290                 295                 300

Glu Leu Lys Leu Lys Glu Ile Gln Leu Gln Glu Arg Glu Arg Ala Leu
305                 310                 315                 320

Lys Ala Arg Glu Glu Arg Leu Glu Gln Lys Glu Gln Glu Leu Cys Val
                325                 330                 335
```

<210> 13
<211> 378
<212> PRT
<213> GAK human aa

<400> 13

```
Met Ser Leu Leu Gln Ser Ala Leu Asp Phe Leu Ala Gly Pro Gly Ser
 1               5                  10                  15

Leu Gly Gly Ala Ser Gly Arg Asp Gln Ser Asp Phe Val Gly Gln Thr
            20                  25                  30

Val Glu Leu Gly Glu Leu Arg Leu Arg Val Arg Arg Val Leu Ala Glu
        35                  40                  45
```

```
Gly Gly Phe Ala Phe Val Tyr Glu Ala Gln Asp Val Gly Ser Gly Arg
    50                  55                  60

Glu Tyr Ala Leu Lys Arg Leu Leu Ser Asn Glu Glu Glu Lys Asn Arg
65                  70                  75                  80

Ala Ile Ile Gln Glu Val Cys Phe Met Lys Lys Leu Ser Gly His Pro
                85                  90                  95

Asn Ile Val Gln Phe Cys Ser Ala Ala Ser Ile Gly Lys Glu Glu Ser
            100                 105                 110

Asp Thr Gly Gln Ala Glu Phe Leu Leu Leu Thr Glu Leu Cys Lys Gly
        115                 120                 125

Gln Leu Val Glu Phe Leu Lys Lys Met Glu Ser Arg Gly Pro Leu Ser
    130                 135                 140

Cys Asp Thr Val Leu Lys Ile Phe Tyr Gln Thr Cys Arg Ala Val Gln
145                 150                 155                 160

His Met His Arg Gln Lys Pro Pro Ile Ile His Arg Asp Leu Lys Val
                165                 170                 175

Glu Asn Leu Leu Leu Ser Asn Gln Gly Thr Ile Lys Leu Cys Asp Phe
            180                 185                 190

Gly Ser Ala Thr Thr Ile Ser His Tyr Pro Asp Tyr Ser Trp Ser Ala
        195                 200                 205

Gln Arg Arg Ala Leu Val Glu Glu Glu Ile Thr Arg Asn Thr Thr Pro
    210                 215                 220

Met Tyr Arg Thr Pro Glu Ile Ile Asp Leu Tyr Ser Asn Phe Pro Ile
225                 230                 235                 240

Gly Glu Lys Gln Asp Ile Trp Ala Leu Gly Cys Ile Leu Tyr Leu Leu
                245                 250                 255

Cys Phe Arg Gln His Pro Phe Glu Asp Gly Ala Lys Leu Arg Ile Val
            260                 265                 270

Asn Gly Lys Tyr Ser Ile Pro Pro His Asp Thr Gln Tyr Thr Val Phe
        275                 280                 285

His Ser Leu Ile Arg Ala Met Leu Gln Val Asn Pro Glu Glu Arg Leu
    290                 295                 300

Ser Ile Ala Glu Val Val His Gln Leu Gln Glu Ile Ala Ala Ala Arg
305                 310                 315                 320

Asn Val Asn Pro Lys Ser Pro Ile Thr Glu Leu Leu Glu Gln Asn Gly
                325                 330                 335
```

Gly Tyr Gly Ser Ala Thr Leu Ser Arg Gly Pro Pro Pro Pro Val Gly
340 345 350

Pro Ala Gly Ser Gly Tyr Ser Gly Gly Leu Ala Leu Ala Glu Tyr Asp
355 360 365

Gln Pro Tyr Gly Gly Phe Leu Asp Ile Leu
370 375

<210> 14
<211> 22
<212> DNA
<213> BUB HRD

<220>
<223> "h" stands for a, c or t.
"y" stands for c or t.
"n" stands for any base.

<400> 14

gaggtgggca thathcaygg na 22

<210> 15
<211> 26
<212> DNA
<213> BUB DVW

<220>
<223> "r" stands for a or g.
"n" stands for any base.
"k" stands for g or t.

<400> 15

gatgaattrt cnggyttnak rtcncc 26

<210> 16
<211> 7
<212> PRT
<213> BUB HRD

<400> 16

Glu Val Gly Ile Ile His Gly
1 5

<210> 17
<211> 8
<212> PRT
<213> BUB DVW

<220>
<223> "Xaa" stands for any base.

<400> 17

Gly Asp Xaa Lys Pro Asp Asn Ser
1 5

<210> 18
<211> 15
<212> PRT
<213> 505A

<400> 18

Gln Lys Tyr Asn Gln Arg Arg Lys His Glu Gln Trp Val Asn Glu
1 5 10 15

<210> 19
<211> 14
<212> PRT
<213> 506A

<400> 19

Asp Asp Lys Asp Glu Trp Gln Ser Leu Asp Gln Asn Glu Asp
1 5 10

<210> 20
<211> 12
<212> PRT
<213> 507A

<400> 20

Lys Asn Ile Gln Lys Phe Val Leu Lys Val Gln Lys
1 5 10

<210> 21
<211> 15
<212> PRT
<213> 508A

<400> 21

Arg Gln Lys Leu Lys Lys Val Phe Gln Gln His Tyr Thr Asn Lys
1 5 10 15

<210> 22
<211> 1058
<212> PRT
<213> BUB1_m

<400> 22

Met Asp Asn Leu Glu Asn Val Phe Arg Met Phe Glu Ala His Met Gln
1 5 10 15

Ser Tyr Thr Gly Asn Asp Pro Leu Gly Glu Trp Glu Ser Phe Ile Lys
20 25 30

Trp Val Glu Glu Asn Phe Pro Asp Asn Lys Glu Tyr Leu Met Thr Leu
35 40 45

Leu Glu His Leu Met Lys Glu Phe Leu His Lys Lys Asn Tyr His Asn
50 55 60

Asp Ser Arg Phe Ile Asn Tyr Cys Leu Lys Phe Ala Glu Tyr Asn Ser
65 70 75 80

Asp Arg His Gln Phe Phe Glu Phe Leu Tyr Asn Gln Gly Ile Gly Thr
85 90 95

Lys Ser Ser Tyr Ile Tyr Met Ser Trp Ala Gly His Leu Glu Ala Gln
100 105 110

Gly Glu Leu Gln His Ala Ser Ala Ile Phe Gln Thr Gly Ile His Asn
115 120 125

Glu Ala Glu Pro Lys Glu Leu Leu Gln Gln Gln Tyr Arg Leu Phe Gln
130 135 140

Ala Arg Leu Thr Gly Ile His Leu Pro Ala Gln Ala Thr Thr Ser Glu
145 150 155 160

Pro Leu His Ser Ala Gln Ile Leu Asn Gln Val Met Met Thr Asn Ser
165 170 175

Ser Pro Glu Lys Asn Ser Ala Cys Val Pro Lys Ser Gln Gly Ser Glu
180 185 190

Cys Ser Gly Val Ala Ser Ser Thr Cys Asp Glu Lys Ser Asn Met Glu
195 200 205

Gln Arg Val Ile Met Ile Ser Lys Ser Glu Cys Ser Val Ser Ser Ser
210 215 220

```
Val Ala Pro Lys Pro Glu Ala Gln Gln Val Met Tyr Cys Lys Glu Lys
225                 230                 235                 240

Leu Ile Arg Gly Asp Ser Glu Phe Ser Phe Glu Glu Leu Arg Ala Gln
                245                 250                 255

Lys Tyr Asn Gln Arg Lys Lys His Glu Gln Trp Val Ser Glu Asp Arg
            260                 265                 270

Asn Tyr Met Lys Arg Lys Glu Ala Asn Ala Phe Glu Glu Gln Leu Leu
        275                 280                 285

Lys Gln Lys Met Asp Glu Leu His Lys Lys Leu His Gln Val Val Glu
    290                 295                 300

Leu Ser His Lys Asp Leu Pro Ala Ser Glu Asn Arg Pro Asp Val Ser
305                 310                 315                 320

Leu Val Cys Val Gly Gln Asn Thr Cys Ser Gln Gln Glu Leu Arg Gly
                325                 330                 335

Pro Ser Leu Ser Ser Ile Ser His Gln Thr Ser Glu Ser Ser Gly Glu
            340                 345                 350

Lys Pro Gln Glu Glu Pro Ser Val Pro Leu Met Val Asn Ala Val Asn
        355                 360                 365

Ser Thr Leu Leu Phe Pro Ala Ala Asn Leu Pro Ala Leu Pro Val Pro
    370                 375                 380

Val Ser Gly Gln Ser Leu Thr Asp Ser Arg Cys Val Asn Gln Ser Val
385                 390                 395                 400

His Glu Phe Met Pro Gln Cys Gly Pro Glu Thr Lys Glu Val Cys Glu
                405                 410                 415

Thr Asn Lys Val Ala Ser Ile Asn Asp Phe His Thr Thr Pro Asn Thr
            420                 425                 430

Ser Leu Gly Met Val Gln Gly Thr Pro Cys Lys Val Gln Pro Ser Pro
        435                 440                 445

Thr Val His Thr Lys Glu Ala Leu Gly Phe Ile Met Asp Met Phe Gln
    450                 455                 460

Ala Pro Thr Leu Pro Asp Ile Ser Asp Asp Lys Asp Glu Trp Pro Ser
465                 470                 475                 480

Leu Asp Gln Asn Glu Asp Ala Phe Glu Ala Gln Phe Gln Lys Asn Ala
                485                 490                 495

Val Ser Ser Gly Asp Trp Gly Val Lys Lys Ile Met Thr Leu Ser Ser
            500                 505                 510
```

Ala Phe Pro Ile Phe Glu Asp Gly Asn Lys Glu Asn Tyr Gly Leu Pro
515 520 525

Gln Pro Lys Asn Lys Pro Leu Gly Ala Arg Thr Phe Gly Glu Arg Ser
530 535 540

Leu Ser Lys Tyr Ser Ser Arg Ser Asn Glu Met Pro His Thr Asp Glu
545 550 555 560

Phe Met Asp Asp Ser Thr Val Cys Gly Ile Arg Cys Asn Lys Thr Leu
565 570 575

Ala Pro Ser Pro Lys Ser Ile Gly Asp Phe Thr Ser Ala Ala Gln Leu
580 585 590

Ser Ser Thr Pro Phe His Lys Phe Pro Ala Asp Leu Val Gln Ile Pro
595 600 605

Glu Asp Lys Glu Asn Val Val Ala Thr Gln Tyr Thr His Met Ala Leu
610 615 620

Asp Ser Cys Lys Glu Asn Ile Val Asp Leu Ser Lys Gly Arg Lys Leu
625 630 635 640

Gly Pro Ile Gln Glu Lys Ile Ser Ala Ser Leu Pro Cys Pro Ser Gln
645 650 655

Pro Ala Thr Gly Gly Leu Phe Thr Gln Glu Ala Val Phe Gly Leu Glu
660 665 670

Ala Phe Lys Cys Thr Gly Ile Asp His Ala Thr Val Glu Asp Leu Ser
675 680 685

Asp Ala Asn Ala Gly Leu Gln Val Glu Cys Val Gln Thr Leu Gly Asn
690 695 700

Val Asn Ala Pro Ser Phe Thr Val Glu Asn Pro Trp Asp Asp Glu Leu
705 710 715 720

Ile Leu Lys Leu Leu Ser Gly Leu Ser Lys Pro Val Thr Ser Tyr Ser
725 730 735

Asn Thr Phe Glu Trp Gln Ser Lys Leu Pro Ala Ile Lys Thr Lys Thr
740 745 750

Glu Tyr Gln Leu Gly Ser Leu Leu Val Tyr Val Asn His Leu Leu Gly
755 760 765

Glu Gly Ala Phe Ala Gln Val Phe Glu Ala Ile His Gly Asp Val Arg
770 775 780

Asn Ala Lys Ser Glu Gln Lys Cys Ile Leu Lys Val Gln Arg Pro Ala
785 790 795 800

Asn Ser Trp Glu Phe Tyr Ile Gly Met Gln Leu Met Glu Arg Leu Lys
805 810 815

Pro Glu Val His His Met Phe Ile Lys Phe Tyr Ser Ala His Leu Phe
820 825 830

Lys Asn Gly Ser Ile Leu Val Gly Glu Leu Tyr Ser Tyr Gly Thr Leu
835 840 845

Leu Asn Val Ile Asn Leu Tyr Lys Asn Thr Ser Glu Lys Val Met Pro
850 855 860

Gln Ala Leu Val Leu Thr Phe Ala Ile Arg Met Leu Tyr Met Val Glu
865 870 875 880

Gln Val His Ser Cys Glu Ile Ile His Gly Asp Ile Lys Pro Asp Asn
885 890 895

Phe Ile Leu Gly His Arg Phe Leu Glu Gln Ala Asp Glu Asp Leu Ala
900 905 910

Thr Gly Leu Ala Leu Ile Asp Leu Gly Gln Ser Ile Asp Met Lys Leu
915 920 925

Phe Pro Lys Gly Thr Val Phe Thr Gly Lys Cys Glu Thr Ser Gly Phe
930 935 940

Gln Cys Pro Glu Met Leu Ser Asn Lys Pro Trp Asn Tyr Gln Ile Asp
945 950 955 960

Tyr Phe Gly Val Ala Ala Thr Ile Tyr Cys Met Leu Phe Gly Ser Tyr
965 970 975

Met Lys Val Lys Asn Glu Gly Gly Val Trp Lys Pro Glu Gly Leu Phe
980 985 990

Arg Arg Leu Pro His Leu Asp Met Trp Glu Glu Phe Phe His Ile Met
995 1000 1005

Leu Asn Ile Pro Asp Cys His Asn Leu Pro Ser Leu Asp Phe Leu Arg
1010 1015 1020

Gln Asn Met Lys Lys Leu Leu Glu Gln Gln Tyr Ser Asn Lys Ile Lys
1025 1030 1035 1040

Thr Leu Arg Asn Arg Leu Ile Val Met Leu Ser Glu Tyr Lys Arg Ser
1045 1050 1055

Arg Lys

<210> 23
<211> 1021
<212> PRT
<213> BUB1_sc

<400> 23

Met Asn Leu Asp Leu Gly Ser Thr Val Arg Gly Tyr Glu Ser Asp Lys
1 5 10 15

Asp Thr Phe Pro Gln Ser Lys Gly Val Ser Ser Ser Gln Lys Glu Gln
20 25 30

His Ser Gln Leu Asn Gln Thr Lys Ile Ala Tyr Glu Gln Arg Leu Leu
35 40 45

Asn Asp Leu Glu Asp Met Asp Asp Pro Leu Asp Leu Phe Leu Asp Tyr
50 55 60

Met Ile Trp Ile Ser Thr Ser Tyr Ile Glu Val Asp Ser Glu Ser Gly
65 70 75 80

Gln Glu Val Leu Arg Ser Thr Met Glu Arg Cys Leu Ile Tyr Ile Gln
85 90 95

Asp Met Glu Thr Tyr Arg Asn Asp Pro Arg Phe Leu Lys Ile Trp Ile
100 105 110

Trp Tyr Ile Asn Leu Phe Leu Ser Asn Asn Phe His Glu Ser Glu Asn
115 120 125

Thr Phe Lys Tyr Met Phe Asn Lys Gly Ile Gly Thr Lys Leu Ser Leu
130 135 140

Phe Tyr Glu Glu Phe Ser Lys Leu Leu Glu Asn Ala Gln Phe Phe Leu
145 150 155 160

Glu Ala Lys Val Leu Leu Glu Leu Gly Ala Glu Asn Asn Cys Arg Pro
165 170 175

Tyr Asn Arg Leu Leu Arg Ser Leu Ser Asn Tyr Glu Asp Arg Leu Arg
180 185 190

Glu Met Asn Ile Val Glu Asn Gln Asn Ser Val Pro Asp Ser Arg Glu
195 200 205

Arg Leu Lys Gly Arg Leu Ile Tyr Arg Thr Ala Pro Phe Phe Ile Arg
210 215 220

Lys Phe Leu Thr Ser Ser Leu Met Thr Asp Asp Lys Glu Asn Arg Ala
225 230 235 240

Asn Leu Asn Ser Asn Val Gly Val Gly Lys Ser Ala Pro Asn Val Tyr
245 250 255

```
Gln Asp Ser Ile Val Val Ala Asp Phe Lys Ser Glu Thr Glu Arg Leu
            260                 265                 270

Asn Leu Asn Ser Ser Lys Gln Pro Ser Asn Gln Arg Leu Lys Asn Gly
        275                 280                 285

Asn Lys Lys Thr Ser Ile Tyr Ala Asp Gln Lys Gln Ser Asn Asn Pro
    290                 295                 300

Val Tyr Lys Leu Ile Asn Thr Pro Gly Arg Lys Pro Glu Arg Ile Val
305                 310                 315                 320

Phe Asn Phe Asn Leu Ile Tyr Pro Glu Asn Asp Glu Glu Phe Asn Thr
                325                 330                 335

Glu Glu Ile Leu Ala Met Ile Lys Gly Leu Tyr Lys Val Gln Arg Arg
            340                 345                 350

Gly Lys Lys His Thr Glu Asp Tyr Thr Ser Asp Lys Asn Arg Lys Lys
        355                 360                 365

Arg Lys Leu Asp Val Leu Val Glu Arg Arg Gln Asp Leu Pro Ser Ser
    370                 375                 380

Gln Pro Pro Val Val Pro Lys Ser Thr Arg Ile Glu Val Phe Lys Asp
385                 390                 395                 400

Asp Asp Asn Pro Ser Gln Ser Thr His His Lys Asn Thr Gln Val Gln
                405                 410                 415

Val Gln Thr Thr Thr Ser Ile Leu Pro Leu Lys Pro Val Val Asp Gly
            420                 425                 430

Asn Leu Ala His Glu Thr Pro Val Lys Pro Ser Leu Thr Ser Asn Ala
        435                 440                 445

Ser Arg Ser Pro Thr Val Thr Ala Phe Ser Lys Asp Ala Ile Asn Glu
    450                 455                 460

Val Phe Ser Met Phe Asn Gln His Tyr Ser Thr Pro Gly Ala Leu Leu
465                 470                 475                 480

Asp Gly Asp Asp Thr Thr Thr Ser Lys Phe Asn Val Phe Glu Asn Phe
                485                 490                 495

Thr Gln Glu Phe Thr Ala Lys Asn Ile Glu Asp Leu Thr Glu Val Lys
            500                 505                 510

Asp Pro Lys Gln Glu Thr Val Ser Gln Gln Thr Thr Ser Thr Asn Glu
        515                 520                 525

Thr Asn Val Arg Tyr Glu Arg Leu Ser Asn Ser Ser Thr Arg Pro Glu
    530                 535                 540
```

```
Lys Ala Asp Tyr Met Thr Pro Ile Lys Glu Thr Thr Glu Thr Asp Val
545                 550                 555                 560

Val Pro Ile Ile Gln Thr Pro Lys Glu Gln Ile Arg Thr Glu Asp Lys
                565                 570                 575

Lys Ser Gly Asp Asn Thr Glu Thr Gln Thr Gln Leu Thr Ser Thr Thr
            580                 585                 590

Ile Gln Ser Ser Pro Phe Leu Thr Gln Pro Glu Pro Gln Ala Glu Lys
        595                 600                 605

Leu Leu Gln Thr Ala Glu His Ser Glu Lys Ser Lys Glu His Tyr Pro
    610                 615                 620

Thr Ile Ile Pro Pro Phe Thr Lys Ile Lys Asn Gln Pro Pro Val Ile
625                 630                 635                 640

Ile Glu Asn Pro Leu Ser Asn Asn Leu Arg Ala Lys Phe Leu Ser Glu
                645                 650                 655

Ile Ser Pro Pro Leu Phe Gln Tyr Asn Thr Phe Tyr Asn Tyr Asn Gln
            660                 665                 670

Glu Leu Lys Met Ser Ser Leu Leu Lys Lys Ile His Arg Val Ser Arg
        675                 680                 685

Asn Glu Asn Lys Asn Pro Ile Val Asp Phe Lys Lys Thr Gly Asp Leu
    690                 695                 700

Tyr Cys Ile Arg Gly Glu Leu Gly Glu Gly Gly Tyr Ala Thr Val Tyr
705                 710                 715                 720

Leu Ala Glu Ser Ser Gln Gly His Leu Arg Ala Leu Lys Val Glu Lys
                725                 730                 735

Pro Ala Ser Val Trp Glu Tyr Tyr Ile Met Ser Gln Val Glu Phe Arg
            740                 745                 750

Leu Arg Lys Ser Thr Ile Leu Lys Ser Ile Ile Asn Ala Ser Ala Leu
        755                 760                 765

His Leu Phe Leu Asp Glu Ser Tyr Leu Val Leu Asn Tyr Ala Ser Gln
    770                 775                 780

Gly Thr Val Leu Asp Leu Ile Asn Leu Gln Arg Glu Lys Ala Ile Asp
785                 790                 795                 800

Gly Asn Gly Ile Met Asp Glu Tyr Leu Cys Met Phe Ile Thr Val Glu
                805                 810                 815

Leu Met Lys Val Leu Glu Lys Ile His Glu Val Gly Ile Ile His Gly
            820                 825                 830
```

```
Asp Leu Lys Pro Asp Asn Cys Met Ile Arg Leu Glu Lys Pro Gly Glu
        835                 840                 845

Pro Leu Gly Ala His Tyr Met Arg Asn Gly Glu Asp Gly Trp Glu Asn
    850                 855                 860

Lys Gly Ile Tyr Leu Ile Asp Phe Gly Arg Ser Phe Asp Met Thr Leu
865                 870                 875                 880

Leu Pro Pro Gly Thr Lys Phe Lys Ser Asn Trp Lys Ala Asp Gln Gln
                885                 890                 895

Asp Cys Trp Glu Met Arg Ala Gly Lys Pro Trp Ser Tyr Glu Ala Asp
            900                 905                 910

Tyr Tyr Gly Leu Ala Gly Val Ile His Ser Met Leu Phe Gly Lys Phe
        915                 920                 925

Ile Glu Thr Ile Gln Leu Gln Asn Gly Arg Cys Lys Leu Lys Asn Pro
    930                 935                 940

Phe Lys Arg Tyr Trp Lys Lys Glu Ile Trp Gly Val Ile Phe Asp Leu
945                 950                 955                 960

Leu Leu Asn Ser Gly Gln Ala Ser Asn Gln Ala Leu Pro Met Thr Glu
                965                 970                 975

Lys Ile Val Glu Ile Arg Asn Leu Ile Glu Ser His Leu Glu Gln His
            980                 985                 990

Ala Glu Asn His Leu Arg Asn Val Ile Leu Ser Ile Glu Glu Glu Leu
        995                 1000                1005

Ser His Phe Gln Tyr Lys Gly Lys Pro Ser Arg Arg Phe
    1010                1015                1020
```

<210> 24
<211> 21
<212> DNA
<213> BUB1 antisense

<400> 24

gtgccaagcc agcagataaa t 21